

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 29, 2004, 10:22:54 ; Search time 18 Seconds  
(without alignments)  
2846.502 Million cell updates/sec

Title: US-10-024-368-2

Perfect score: 5358

Sequence: 1 FYYKDFSIPLAKEPLAAAE.....MDNAVFQAVKAVGTQLEAK 984

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_42:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	5350	99.9	1189	1	HAIR_HUMAN
2	4110.5	76.7	1182	1	HAIR_MOUSE
3	4067	75.9	1181	1	HAIR_RAT
4	694.5	13.0	1214	1	TSGA_RAT
5	205	3.8	1616	1	APXL_HUMAN
6	190.5	3.6	660	1	YHL1_EBV
7	189	3.5	2167	1	SHK1_RAT
8	186.5	3.5	1461	1	IE18_PRIVF
9	183	3.4	1446	1	IE18_PRIVKA
10	182.5	3.4	1723	1	KA93_HUMAN
11	180.5	3.4	2161	1	SHK3_HUMAN
12	180	3.4	1733	1	VNUA_PRIVKA
13	176	3.3	2774	1	MAPA_RAT
14	173	3.2	1783	1	RAA3_CHLRE
15	172	3.2	1670	1	CA34_HUMAN
16	171.5	3.2	2142	1	BAT2_HUMAN
17	170.5	3.2	5262	1	MLL2_HUMAN
18	168.5	3.1	1466	1	CA13_HUMAN
19	167.5	3.1	1017	1	EM11_MOUSE
20	166.5	3.1	1143	1	CA11_HUMAN
21	166	3.1	1372	1	CA21_MOUSE
22	164	3.1	1464	1	CA13_MOUSE
23	163.5	3.1	1411	1	TCOF_HUMAN
24	163.5	3.1	1459	1	CA12_MOUSE
25	163	3.0	1366	1	CA21_HUMAN
26	162	3.0	1690	1	CA44_HUMAN
27	162	3.0	1691	1	CA64_HUMAN
28	162	3.0	2944	1	CA17_HUMAN
29	161.5	3.0	1603	1	CA1F_HUMAN
30	160	3.0	1262	1	CA13_CHICK
31	160	3.0	1366	1	CA21_CANFA
32	159	3.0	1372	1	CA21_RAT
33	159	3.0	1418	1	CA12_HUMAN

## RESULT 1

ID	HAIR_HUMAN	STANDARD;	PRT;	1189 AA.
AC	O43593; Q96H33; Q9NPE1;			
DT	15-JUL-1999 (Rel. 38, Created)			
DT	10-OCT-2003 (Rel. 42, Last sequence update)			
DT	10-OCT-2003 (Rel. 42, Last annotation update)			
DE	Hairless protein.			
GN	HR.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A. (ISOFORM 1), AND VARIANT ALUNC ALA-1022.			
RX	MEDLINE=98111413; PubMed=9445480;			
RA	Almad W., ul Haque M.F., Brancolini V., Tsou H.C., Ul Haque S., Lam H., Aita V.M., Owen J., Deblaquiere M., Frank J., Cserhalmi-Friedman P.B., Leask A., McGrath J.A., Peacocke M., Almad M., Ott J., Christiano A.M.;			
RA	"Alopecia universalis associated with a mutation in the human hairless gene.";			
RT	Science 279:720-724(1998).			
RL	[2]			
RN	SEQUENCE FROM N.A. (ISOFORM 1), REVISIONS TO 572 AND 774, AND TISSUE SPECIFICITY.			
RP	TISSUE=peripheral blood leukocytes, and Skin fibroblast;			
RX	MEDLINE=99162400; PubMed=10051399;			
RA	Almad W., Zlotogorski A., Panteleyev A.A., Lam H., Ahmad M., ul Haque M.F., Abdallah H.M., Dragan L., Christiano A.M.;			
RA	"Genomic organization of the human hairless gene (HR) and identification of a mutation underlying congenital atrichia in an Arab Palestinian family.";			
RT	Genomics 56:141-148(1999).			
RL	[3]			
RN	REVISIONS TO 446 AND 584.			
RP	Christiano A.M.;			
RP	Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.			
RL	[4]			
RN	SEQUENCE FROM N.A. (ISOFORMS 1 AND 2), VARIANT ALUNC ASP-1136, AND TISSUE SPECIFICITY.			
RP	TISSUE=Brain, Fetal brain, and Peripheral blood leukocytes;			
RX	MEDLINE=98409496; PubMed=9736769;			
RA	Cichon S., Anker M., Vogt I.R., Rohleder H., Putzstuck M., Hillmer A., Farooq S.A., Al-Dhafri K.S., Ahmad M., Haque S., Rietschel M., Propping P., Kruse R., Noethen M.M.;			
RA	"Cloning, genomic organization, alternative transcripts and mutational analysis of the gene responsible for autosomal recessive universal congenital alopecia.";			
RT	Hum. Mol. Genet. 7:1671-1679(1998).			
RL	[5]			
RN	SEQUENCE OF 198-1189 FROM N.A. (ISOFORM 3).			
RP	TISSUE=Skin;			
RX	MEDLINE=22388257; PubMed=12477932;			
RA	Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G., Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,			

RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
RA Diachenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,  
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,  
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
RA Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,  
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U.,  
RA Scherch A., Schein J.E., Jones S.J.M., Marra M.A.,  
RT "Generation and initial analysis of more than 15,000 full-length human  
RT and mouse cDNA sequences.";  
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
RN [6]  
RP VARIANT GLN-620.  
RX MEDLINE=98431781; PubMed=9758627;  
RA Ahmad W., Irvine A.D., Lam H., Buckley C., Bingham B.A.,  
RA Panteleyev A.A., Ahmad M., McGrath J.A., Christiano A.M.;  
RT "A missense mutation in the zinc-finger domain of the human hairless  
RT gene underlies congenital atrichia in a family of Irish travellers.";  
RL Am. J. Hum. Genet. 63:984-991(1998).  
RN [7]  
RP VARIANT GLN-620.  
RX MEDLINE=21303039; PubMed=11410842;  
RA Hillmer A.M., Kruse R., Betz R.C., Schumacher J., Heyn U.,  
RA Propping P., Noethen M.M., Cichon S.;  
RT "Variant 1859G-->A (Arg620Gln) of the 'hairless' gene: absence of  
RT association with papular atrichia or androgenic alopecia.";  
RL Am. J. Hum. Genet. 69:235-237(2001).  
RN [8]  
RP VARIANT ALUNC-1012.  
RX MEDLINE=22294554; PubMed=12406339;  
RA Klein I., Bergman R., Indelman M., Sprecher E.;  
RT "A novel missense mutation affecting the human hairless thyroid  
RT receptor interacting domain 2 causes congenital atrichia.";  
RL J. Invest. Dermatol. 119:920-922(2002).  
CC -!- FUNCTION: May act as a transcription factor that could act on to  
CC regulate one of the phases of hair growth.  
CC -!- SUBCELLULAR LOCATION: Nuclear.  
CC -!- ALTERNATIVE PRODUCTS:  
CC Event=Alternative splicing; Named isoforms=3;  
CC Comment=Additional isoforms seem to exist;  
CC Name=1; Synonyms=Long;  
CC IsoId=O43593-1; Sequence=Displayed;  
CC Name=2; Synonyms=Short;  
CC IsoId=O43593-2; Sequence=VSP\_004276;  
CC Name=3;  
CC IsoId=O43593-3; Sequence=VSP\_007990, VSP\_007991;  
CC Note=No experimental confirmation available;  
CC -!- TISSUE SPECIFICITY: Strongest expression of isoforms 1 and 2 is  
CC seen in the small intestine, weaker expression in brain and colon,  
CC and trace expression is found in liver, pancreas, spleen, thymus,  
CC stomach, salivary gland, appendix and trachea. Isoform 1 is always  
CC the most abundant. Isoform 1 is exclusively expressed at low  
CC levels in kidney and testis and isoform 2 exclusively at high  
CC levels in the skin.  
CC -!- DISEASE: Defects in HR are the cause of alopecia universalis  
CC congenita (ALUNC) [MIM:203655]. ALUNC is a rare autosomal  
CC recessive form of hair loss characterized by hair follicles  
CC without hair.  
CC -!- DISEASE: Defects in HR are the cause of atrichia with papular  
CC lesions (APL) [MIM:209500]; also known as congenital atrichia. APL  
CC is an autosomal recessive disease characterized by papillary  
CC lesions over most of the body and almost complete absence of hair.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL: AF039196; AAC32258.3; -.  
DR EMBL: AJ277249; CAB87577.2; -.  
DR EMBL: AJ277250; CAB87577.2; JOINED.  
DR EMBL: AJ277251; CAB87577.2; JOINED.  
DR EMBL: AJ277252; CAB87577.2; JOINED.  
DR EMBL: AJ277253; CAB87577.2; JOINED.  
DR EMBL: AJ400825; CAB87577.2; JOINED.  
DR EMBL: AJ400826; CAB87577.2; JOINED.  
DR EMBL: AJ400827; CAB87577.2; JOINED.  
DR EMBL: AJ400828; CAB87577.2; JOINED.  
DR EMBL: AJ400829; CAB87577.2; JOINED.  
DR EMBL: AJ400830; CAB87577.2; JOINED.  
DR EMBL: AJ400831; CAB87577.2; JOINED.  
DR EMBL: AJ400832; CAB87577.2; JOINED.  
DR EMBL: AJ400833; CAB87577.2; JOINED.  
DR EMBL: AJ400834; CAB87577.2; JOINED.  
DR EMBL: AJ400835; CAB87577.2; JOINED.  
DR EMBL: AJ400836; CAB87577.2; JOINED.  
DR EMBL: AJ277165; CAB87577.2; JOINED.  
DR EMBL: BC008946; AAB08946.1; -.  
DR EMBL: BC008946; AAB08946.1; ALT\_INIT.  
DR Genew; HGNC:5172; HR.  
DR MIM; 602302; -.  
DR MIM; 203655; -.  
DR MIM; 209500; -.  
DR GO; GO:0003634; C:nucleus; NAS.  
DR GO; GO:0003700; F:transcription factor activity; NAS.  
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; NAS.  
DR InterPro; IPR003347; TF\_JmJC.  
DR Pfam; PF02373; jmjC; 1.  
DR SMART; SM00558; jmjC; 1.  
KW Zinc-finger; DNA-binding; Nuclear protein; Transcription regulation;  
KW Metal-binding; Alternative splicing; Disease mutation; Polymorphism.  
FT ZN FING 600 625  
FT VARSPLIC 979 994 LALRPLEPOLWAAVGV -> AGPGEHSQRHSALPLP (in  
FT isoform 3).007990.  
FT /FTId=VSP\_007990.  
FT Missing (in isoform 3).  
FT /FTId=VSP\_007991.  
FT Missing (in isoform 2).  
FT /FTId=VSP\_004276.  
FT R -> O.  
FT /FTId=VAR\_005265.  
FT D -> N (in ALUNC); affects binding to  
FT thyroid hormone receptor).  
FT /FTId=VAR\_016222.  
FT T -> A (in ALUNC).  
FT /FTId=VAR\_005266.  
FT V -> D (in ALUNC).  
FT /FTId=VAR\_005267.  
FT D -> G (IN REF. 4).  
FT S -> L (IN REF. 5).  
SQ SEQUENCE 1189 AA; 127552 MW; B4DC9C6292AA2018 CRC64;  
Query Match 99.9%; Score 5350; DB 1; Length 1189;  
Best Local Similarity 99.9%; Pred. No. 3.9e-286;  
Matches 983; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 FFKDPSIPRLAKEPLAAAEPLGLNSGHLORAGEAERPSLHQRDGEMGAGRQQNCP 60  
DB 206 FFKDPSIPRLAKEPLAAAEPLGLNSGHLORAGEAERPSLHQRDGEMGAGRQQNCP 265  
QY 61 LFLGQDPTVPTWSPACPPGLVHTLGNVWAGPDGNLGYOLGPPATPCPSPPPTVQRG 120  
DB 266 LFLGQDPTVPTWSPACPPGLVHTLGNVWAGPDGNLGYOLGPPATPCPSPPPTVQRG 325  
QY 121 CCSYSPYTKGDLGPCGCKOEGLEGASGASEPSEVNVKASGPACPPSHHTKLKTLWT 180

dry Match 76.7%; Score 4110.5; DB 1; Length 1182;  
t Local Similarity 78.5%; Pred. No. 3.5e-218;  
ches 772; Conservative 62; Mismatches 143; Indels 7; Gaps 5;  
1 FYYKDPISPLAKEPLAABPGFLGNSGGHLOKAGEAERPSLHORDGEMAGRONPCP 60

RN [1]



QY 481 IRGHCPCQADARVWAPGDAGQOKESTOKTPTTPOPCSCNGDTHRTKSIKEETPDPAETPAE 540  
 DB 680 IRGHCFCQDARVWAPGDGQOKEPTKTPPAQLSCNGSDNRTKDIKEETPDSTESPAP 739  
 QY 541 DRAGRGLPCPSICELASTAVKLCIGHERIHMAFAPVTPALPSDDRTITILDSIIAQVV 600  
 DB 740 DRAGRSPLPCPSICELASTAVKLCIGHERIHMAFAPVTPALPSDDRTITILDSIIAQVV 739  
 QY 601 ERKIQEKALGFLRAGPLRGKGLPLSPVRPRLPPFGALLWLOEQPCPRRGFHLFQEH 660  
 DB 800 ERKIQEKALGFLRAGSLRGKGLPLSPVRTQLSPFGALLWLOEPR-PKHGFLFQEH 857  
 QY 661 WRQOPVLVSGIORTLOGLNWTEALGALGOVALSPGLPPQSPSSLGSTTFWEGFSWPE 720  
 DB 858 WRQOPVLVSGIORTLRLSLWGMELTGLGQVOTTLTALGPPQPTSLDSTAFWKGFSPHE 917  
 QY 721 LRPKSDGSLVLLHRLALGDEDTSRVENLAASLPLPEYCALHGKLNLASYLPPGGLALRPLE 780  
 DB 918 ARPCLDGSVLLHRLPLGDXDSRVENLAASLPLPEYCAHQKLNLASYLPLGLTLHLE 977  
 QY 781 POLMAAYGVSPHRGLGTKNLCVEVADLVSLVHADTLPFAWHRAQKDFLSGLDGEGLWS 840  
 DB 978 POLMAAYGVNSHRGLGTKNLCVEVADLVSLVHADTLPFAWHRAQKDFLSGLDGEGLWS 1037  
 QY 841 PGSOVSTVWVFRADQRIIRRFLOWYCPAGAGALEPGACGSCYLDAGLRRLREEWGVS 900  
 DB 1038 PGSOVSTVWVFRADQRIIRRFLOWYCPAGAGALEPGACGSCYLDAGLRRLREEWGVS 1097  
 QY 901 CWTLLQAPGEAVLPAGAPHQVQGLVSTVSTVTHFLSPETSALSQALCHQGPSLPPDCHL 960  
 DB 1098 CWTLLQAPGEAVLPAGAPHQVQGLVSTVSTVTHFLSPETSALSQALCHQGPSLPPDCHL 1157  
 QY 961 LYAQMMAVFOAVKAVGTLOEAK 984  
 DB 1158 LYAQMMAVFOAVKAVGTLOEAK 1181

RESULT 4  
 TSQA RAT STANDARD; PRT; 1214 AA.  
 AC Q63679;  
 DT 15-JUL-1998 (Rel. 36, Created)  
 DT 15-JUL-1998 (Rel. 36, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE Testis specific protein A (Zinc finger protein TSQA).  
 GN TSQA.  
 OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.  
 OX NCBI\_TaxID=10116;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Sprague-Dawley; TISSUE=Testis;  
 RX MEDLINE=92172411; PubMed=1793593;  
 RA Hoog C., Schalling M., Brundell E., Daneholt B.;  
 RT "Analysis of a murine male germ cell-specific transcript that encodes  
 a putative zinc finger protein."  
 RL Mol. Reprod. Dev. 30:173-181(1991).  
 CC -!- FUNCTION: Meiotic or postmeiotic function.  
 CC -!- TISSUE SPECIFICITY: Testis specific. Expressed only in male germ  
 cells.  
 CC -!- DEVELOPMENTAL STAGE: Reaches a maximum during the meiotic and the  
 postmeiotic stages of germ cell development.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See [http://www.isb-sib.ch/](http://www.isb-sib.ch/announce/)  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; X59993; CAN42610.1; -.

DR PIR; S28499; S28499.  
 DR InterPro; IPR007113; Cupin sup.  
 DR InterPro; IPR003347; TF\_JmJC.  
 DR Pfam; PF02373; JmJC; 1.  
 DR SMART; SM00558; JmJC; 1.  
 KW Zinc-finger; Metal-binding.  
 FT ZN FING 546 571  
 SQ SSEQUENCE 1214 AA; 135403 MW; 86552220760EA6E CRC64;  
 Query Match 13.0%; Score 694.5; DB 1; Length 1214;  
 Best Local Similarity 25.8%; Pred. No. 7.2e-31;  
 Matches 260; Conservative 159; Mismatches 397; Indels 192; Gaps 40;  
 QY 113 EPPVTOGCCSSPPYTKG-GDL-----GPCKQCEGLEGGAGSEPEEVNKAAGPRACP 167  
 DB 244 KPDVCKAGLLSSKSSQVAGDLKILSEPKGSC-----IQPKTNDQESRLSESTP 292  
 QY 168 PSHTTKLKTWLRHRSQFECPRG-CPEVERPVARLRALKRAGS-----PEVQGA- 217  
 DB 293 QPVTGLTKELVTKTSSKAELDNATAPLOKR-----LEHTASTPDGLSDKEVEAGV 345  
 QY 218 --MGSPAPKPPDPFPPTAGAGGQVEVDTSTGNKVDVS---GQHDEQKPGDQOQASL 272  
 DB 346 TRLNSCSEKK---VGPSDLGSQS-QNLKETS-VKDHDSCTTRSSNKTQTPPARKSVL 398  
 QY 273 QDPG-----LQDIPCLALPAKLAQCSCA-----QAAGEGGHACHSQOVR 313  
 DB 399 TDPDKLKLQSGEAFVDDSCVNIQAQLPKRECLDSLRKDKQKQDSVPVCRFFHER 458  
 QY 314 RSLPG--GELQOEBDTATNSSSEEGPG-----SGPDSRLSTGLAKHLLSLGLDRLC 362  
 DB 459 RLQENKHGVLVEGFLTPNKYDSEALGLWLP LTKNVVGTD--LDT--AKYILANIGDHFC 514  
 QY 363 RLLREREAL-----AWAQREGGPAVTDSPGIPRCCSCCHHGLFNTHWRCPKCS 413  
 DB 515 QMWISEKEAMSTIEPHRQVAVKR-----AV---KGVREMCVDCDTTIFNLHWVCPRCG 564  
 QY 414 HRLVACGRVAGTGRAREKAGFQESAECTQEAAGAACSLMLTFVSSQALAEIATMH 473  
 DB 565 FGVCVDCYRLK-RKNCQOQAAKYTFSWIRCVKQSIHEFENLMPQTQIPGKALYDVGDIVH 623  
 QY 474 QVWVKFDIRGHCPQADARVWAPGDAGQOKESTOKTPTTPOPCSCNGDTHRTKSIKEETPD 533  
 DB 624 SVRAKWKIKNPCPC-SNRQFKLFSKPAKEDLKASLSGEKPSLGTWVQOSSPVLEPAV 682  
 QY 534 SAETPAEDRAGRGLPCPS-----LCELLASTAVK-----LCLGHERIHMAFAPVTP 580  
 DB 683 CGEAPSPASNVKPI-CPANTSPNLWADLTSGNVNKENKEKQLTMPILKNEIKCLPPLP 741  
 QY 581 ALPSDDRTITILDSIIAQVVE-----RKIQEALGP---GLRAGPCL----- 619  
 DB 742 PLNKSSTVLTFTFNSTILTTPVNNNSGFLRNJLNSSTGTENGLKNTPKILDDIFASLVQN 801  
 QY 620 -----RKGLGLPLSPVRPRLPPP-----GALLWLOEQPCPRRGFHLFQEHWRQOQP 666  
 DB 802 KTSDDLKRPQGLTIKPSILGFDTPHYWLCNRLCLQDPN--NKSNNVWFECWKQOQP 859  
 QY 667 VLVSGIORTLOGLNWTEALGALGOVALSPGLPPQSPSSLGSTT--FWEGFSWPELRPK 724  
 DB 860 VVWVGHHKLNTELWKPESFRKEFGEQEVLDVNCRTNEIITGATVGDVFDGFDVDPNRLK 919  
 QY 725 SD---ESVLLHRLALGDED-----TSRVENLAASLPLPEYCALHGKLNLASYLPPGLAL 776  
 DB 920 NEKEPMLVKLQDWPGEDFRDMMPSRFDLMANILPLPEYTRDRDGLKLNLASRLPNVFR 979  
 QY 777 RPLPQLWAAVY-VSPHRGHLGTKNLCVEVADLVSLVHADTLPFAWHRAQKDFLSGLDG 835  
 DB 980 PDLGPKMAYGLTTPEDRKYGTTNLHLDSDAANVWVYVIGIPKQCEQEVEVLTIQDG 1039  
 QY 836 -----EGLWSPGQSQVSTVWVFRADQRIIRRFLOWYCPAGAGALEPGACGSCYLDAGLRRLREEWGVS 877  
 DB 1040 DSELTIKRFTIEGKEFGA---LWHYAAKDTKEIKREFLKKVSEEQEQENPADH---DP 1092

```

QY      878  GAGSCVLDAGLRRLREEWGVCWTLLOAPGEANVLVPAGAPHQVQGLVSTVSTQHFLS 9337
Db      1093  IHDQSWYLDRLSKRLLYGXEYQGWAIQVQLGVDVFFIPAGAPHQVHNLVSCIKVAEDFVS 1152
QY      938  PETSAS---LSAQLCHQGPSLPPDCHLLYAQMDWAVQAVKVAVGTILQ 981
Db      1153  PEHVKHCFWLTQEFRLH--SQTHNHEDKLQKNVIYHAKVDAMLUK 1198

RESULT 5
APXL HUMAN STANDARD; PRT; 1616 AA.
AC Q1376;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE DE Apical-like protein (APXL protein).
GN APXL.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OC NCBI_TaxId=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Retina;
RX MEDLINE=95315933, PubMed=7795590;
RA Schiaffino V.M., Bassi M.T., Rugarli E.I., Renieri A., Galli L.,
RA Ballabio A.;
RT "Cloning of a human homologue of the Xenopus laevis APX gene from the
RT ocular albinism type 1 critical region.";
RL Hum. Mol. Genet. 4:373-382(1995).
RN [2]
RP SEQUENCE OF 56-1616 FROM N.A.
RA Shen Y., Gibbs R.A.;
RL Submitted (JAN-1997) to the EMBL/GenBank/DBJ databases.
CC -1- TISSUE SPECIFICITY: Abundant in retina and melanoma; also in
CC brain, placenta, lung, kidney and pancreas.
CC -1- SIMILARITY: Contains 1 PDZ/DHR domain.
CC -1- SIMILARITY: SOME, TO XENOPUS LAEVIS APICAL PROTEIN (APX).
CC ---
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC ---
DR EMBL; X83543; CAA58534.1; --
DR EMBL; AC002365; AAC32592.1; --
DR PIR; I37183; I37183.
DR HSPSP; P29476; IQAV.
DR Genew; HGNC:630; APXL.
DR MIM; 300103; --
DR GO; GO:0015280; F:amiloride-sensitive sodium channel activity; TAS.
DR InterPro; IPR001478; PDZ.
DR Pfam; PF005595; PDZ; 1.
DR SMART; SMO0228; PDZ; 1.
DR PROSITE; PSS0106; PDZ; 1.
DR DOMAIN 26 108 PDZ.
DR DOMAIN 150 153 POLY-SER.
DR DOMAIN 314 320 POLY-PRO.
DR DOMAIN 343 346 POLY-ALA.
DR DOMAIN 1065 1068 POLY-PRO.
SQ SEQUENCE 1616 AA; 176409 MW; 752406B58C0B60A2 CRC64;

```

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 29, 2004, 11:48:34 ; Search time 12295 Seconds  
(without alignments)  
11287.873 Million cell updates/sec

Title: US-10-024-368-1  
Perfect score: 3202  
Sequence: 1 ttttactacaagatccgag.....tgtgccttttatgtacagg 3202

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 3470272 seqs, 21671516995 residues  
Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : GenEmbl.\*

- 1: gb.ba.\*
- 2: gb.htg.\*
- 3: gb.in.\*
- 4: gb.om.\*
- 5: gb.ov.\*
- 6: gb.pat.\*
- 7: gb.ph.\*
- 8: gb.pl.\*
- 9: gb.pr.\*
- 10: gb.ro.\*
- 11: gb.sts.\*
- 12: gb.sy.\*
- 13: gb.un.\*
- 14: gb.vi.\*
- 15: em.ba.\*
- 16: em.fun.\*
- 17: em.hum.\*
- 18: em.in.\*
- 19: em.mu.\*
- 20: em.om.\*
- 21: em.or.\*
- 22: em.ov.\*
- 23: em.pat.\*
- 24: em.ph.\*
- 25: em.pl.\*
- 26: em.ro.\*
- 27: em.sts.\*
- 28: em.un.\*
- 29: em.vi.\*
- 30: em.htg.hum.\*
- 31: em.htg.inv.\*
- 32: em.htg.other.\*
- 33: em.htg.mus.\*
- 34: em.htg.pln.\*
- 35: em.htg.rod.\*
- 36: em.htg.mam.\*
- 37: em.htg.vrt.\*
- 38: em.sy.\*
- 39: em.htgo.hum.\*
- 40: em.htgo.mus.\*
- 41: em.htgo.other.\*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	3202	100.0	3202	6	AR194210	Sequence
2	3185	99.5	5709	9	AF039196	AF039196 Homo sapi
3	3042.4	95.0	3719	9	HS277165	AJ277165 Homo sapi
4	2817.4	88.0	3564	9	AF361864	AF361864 Macaca mu
5	2308.4	72.1	3210	9	BC008946	BC008946 Homo sapi
6	1985.4	62.0	5599	10	BC049182	BC049182 Mus muscu
7	1970.8	61.5	5186	10	RNU71293	U71293 Rattus norv
8	1970	61.5	4017	10	MMHAIPRO	Z32675 M.musculus
9	1963.6	61.3	4017	10	AY130969	AY130969 Mus muscu
10	943.4	29.5	2128	4	AY130969	AY130969 Ovis arie
11	828	25.9	1027	4	AY279972	AY279972 Sus scrof
12	786.8	24.6	218931	9	AC105206	AC105206 Homo sapi
13	786.2	24.6	958	9	HS277250	AJ277250 Homo sapi
14	595.8	18.6	2394	6	AX835172	AX835172 Sequence
15	595.8	18.6	2394	9	AK098053	AK098053 Homo sapi
16	533.2	16.7	538	6	AX677839	AX677839 Sequence
17	460.2	14.4	207136	10	AC122268	AC122268 Mus muscu
18	460.2	14.4	219288	2	AC131321	AC131321 Mus muscu
19	449.4	14.0	255506	2	AC097410	AC097410 Rattus no
20	338.8	10.6	45185	2	AC011166	AC011166 Homo sapi
21	306	9.6	3357	9	HSM804566	AL833253 Homo sapi
22	279.8	8.7	231632	2	AC119127	AC119127 Rattus no
23	258.5	8.1	43402	2	AC105207	AC105207 Homo sapi
24	243	7.6	448	9	HS277252	AJ277252 Homo sapi
25	197.8	6.2	582	9	HS277252	AJ277252 Homo sapi
26	177.6	5.5	345	9	HS277252	AJ277252 Homo sapi
27	171.4	5.4	454	9	HS277252	AJ277252 Homo sapi
28	171.4	5.4	43402	2	AC105207	AC105207 Homo sapi
29	169.4	5.3	542	9	HS277253	AJ277253 Homo sapi
30	168	5.2	691	9	HS277253	AJ277253 Homo sapi
31	166	5.2	458	9	HS277253	AJ277253 Homo sapi
32	150.2	4.7	376	9	HS277251	AJ277251 Homo sapi
33	140.2	4.4	410	9	HS277251	AJ277251 Homo sapi
34	133.4	4.2	464	9	HS277251	AJ277251 Homo sapi
35	133.4	4.2	45185	2	AC011166	AC011166 Homo sapi
36	122.2	3.8	516	9	HS277253	AJ277253 Homo sapi
37	121.2	3.8	367	9	HS277253	AJ277253 Homo sapi
38	117.8	3.7	305	9	HS277253	AJ277253 Homo sapi
39	103.6	3.2	419	10	MMMS39	AJ002245 Mus muscu
40	98.4	3.1	309	9	HS277253	AJ277253 Homo sapi
41	86.6	2.7	125020	9	AF429315	AF429315 Homo sapi
42	85.8	2.7	301	11	BV014248	BV014248 S212P6107
43	85.8	2.7	490	9	HS277253	AJ277253 Homo sapi
44	76	2.4	125020	9	AF429315	AF429315 Homo sapi
45	73.2	2.3	376	9	HS277253	AJ277253 Homo sapi

ALIGNMENTS

RESULT 1  
AR194210  
LOCUS AR194210 3202 bp DNA linear PAT 20-APR-2002  
DEFINITION Sequence 1 from patent US 6348348.  
ACCESSION AR194210  
VERSION AR194210.1 GI:20240802  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 3202)  
AUTHORS Thompson,C.C.  
TITLE Human hairless gene and protein  
JOURNAL Patent: US 6348348-A 1 19-FEB-2002;  
FEATURES Location/Qualifiers

source	1..3202				
ORIGIN	/organism="unknown"				
	/mol_type="unassigned DNA"				
Query Match	100.0%;	Score 3202;	DB 6;	Length 3202;	
Best Local Similarity	100.0%;	Pred. No. 0;			
Matches 3202;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;	
QY	1	TTTTACTACAGGATCCGAGCATTTCCAGGTTGGCAAGGAGCCCTTGGCAGCTGGCGAA	60		
Db	1	TTTTACTACAGGATCCGAGCATTTCCAGGTTGGCAAGGAGCCCTTGGCAGCTGGCGAA	60		
QY	61	CTTGGGTTGTTGGCTTAAACTCTGTGGGCACCTGCAGAGAGCCGGGGAGGCCGAAACGC	120		
Db	61	CTTGGGTTGTTGGCTTAAACTCTGTGGGCACCTGCAGAGAGCCGGGGAGGCCGAAACGC	120		
QY	121	CTTCTACTGACACAGAGGATGGAGATGGGAGCTGGCCGGCAGCAGAAATCCTTGGCCCG	180		
Db	121	CTTCTACTGACACAGAGGATGGAGATGGGAGCTGGCCGGCAGCAGAAATCCTTGGCCCG	180		
QY	181	CTCTTCTGGGGCAGCAGACACTGTGCCCTGGACCTCTCTGGCCCGCTTGTTCCTCCAGGC	240		
Db	181	CTCTTCTGGGGCAGCAGACACTGTGCCCTGGACCTCTCTGGCCCGCTTGTTCCTCCAGGC	240		
QY	241	CTTGTTCATATCTTGGCAAACGTTCTGGGCTGGGCCAGGGCATGGGAACCTTGGGTACCAG	300		
Db	241	CTTGTTCATATCTTGGCAAACGTTCTGGGCTGGGCCAGGGCATGGGAACCTTGGGTACCAG	300		
QY	301	CTGGGGCCACCAAGCAACCAAGGTGGCCCTCTCTCTGAGCCGCTGTCAACAGCGGGGC	360		
Db	301	CTGGGGCCACCAAGCAACCAAGGTGGCCCTCTCTCTGAGCCGCTGTCAACAGCGGGGC	360		
QY	361	TGCTGTTCATCTTACCCACCACTAAAGTGGGGATCTTGGCCCTTGTGGGAAGTGCCAG	420		
Db	361	TGCTGTTCATCTTACCCACCACTAAAGTGGGGATCTTGGCCCTTGTGGGAAGTGCCAG	420		
QY	421	GAGGGCTCGAGGGGGTGCAGTGGAGCCAGCGAACCCAGCAGGAAGTGAAACAAGGCC	480		
Db	421	GAGGGCTCGAGGGGGTGCAGTGGAGCCAGCGAACCCAGCAGGAAGTGAAACAAGGCC	480		
QY	481	TCTGGCCCCAGGGCTGTCCCCCAGCCACCAACCAAGCTGAAGACATGGCTACA	540		
Db	481	TCTGGCCCCAGGGCTGTCCCCCAGCCACCAACCAAGCTGAAGACATGGCTACA	540		
QY	541	CGGCACTCGAGGACGTTTGAATGTCCACGGGCTGCCCTCAGGTTCAGAGAGAGCCGGTT	600		
Db	541	CGGCACTCGAGGACGTTTGAATGTCCACGGGCTGCCCTCAGGTTCAGAGAGAGCCGGTT	600		
QY	601	GCTCGGCTCGGGCCCTCAAAAGGGCAGGAGCCCGAGGTCAGGGAGCAATGGGCAGT	660		
Db	601	GCTCGGCTCGGGCCCTCAAAAGGGCAGGAGCCCGAGGTCAGGGAGCAATGGGCAGT	660		
QY	661	CCAGCCCCCAAGGGCCACCGGACCTTTCCAGGCACTGCAGAACAGGGGGTGGGGGT	720		
Db	661	CCAGCCCCCAAGGGCCACCGGACCTTTCCAGGCACTGCAGAACAGGGGGTGGGGGT	720		
QY	721	TGGCAGAGGTTTCGGGACACATCGATAGGAAACAAGATGTGGACTCGGGACAGCATGAT	780		
Db	721	TGGCAGAGGTTTCGGGACACATCGATAGGAAACAAGATGTGGACTCGGGACAGCATGAT	780		
QY	781	GAGCAGAAAGGACCCCAAGATGCGCAGGCAAGTCTCAGAACCCGGGACTTCAGGACATA	840		
Db	781	GAGCAGAAAGGACCCCAAGATGCGCAGGCAAGTCTCAGAACCCGGGACTTCAGGACATA	840		
QY	841	CCATGCTCTTCTCCCTGCAAACTGGCTCAATGCCAAAGTTGTGCCAGGCAGCTGGA	900		
Db	841	CCATGCTCTTCTCCCTGCAAACTGGCTCAATGCCAAAGTTGTGCCAGGCAGCTGGA	900		
QY	901	GAGGAGGAGGGCAGCCTGCTCCACTCTCAGCAAGTCCGAGATTCGCTCTGGGAGGGGAG	960		
Db	901	GAGGAGGAGGGCAGCCTGCTCCACTCTCAGCAAGTCCGAGATTCGCTCTGGGAGGGGAG	960		
QY	961	CTGACAGAGGAGGACACACAGCCACCACCTCCAGCTCTGAGGAAGGCCCCAGGGTCCGGC	1020		
Db	961	CTGACAGAGGAGGAGGACACACAGCCACCACCTCCAGCTCTGAGGAAGGCCCCAGGGTCCGGC	1020		
QY	1021	CTTGACAGCCGGCTCAGCACAGGCTCTGCCAAGCACCTGCTCAGTGTGTTGGGGAGCCGA	1080		
Db	1021	CTTGACAGCCGGCTCAGCACAGGCTCTGCCAAGCACCTGCTCAGTGTGTTGGGGAGCCGA	1080		
QY	1081	CTGTGCCGCTGTCTGGGAGGAGCGGAGGCCCTTGGTTCGGCCACAGCGGGAAGGCCAA	1140		
Db	1081	CTGTGCCGCTGTCTGGGAGGAGCGGAGGCCCTTGGTTCGGCCACAGCGGGAAGGCCAA	1140		
QY	1141	GGGCCAGCCGTGACAGGGGACAGCCAGGCAATTCACGCTGTCTGACGCGTTCCCAACAT	1200		
Db	1141	GGGCCAGCCGTGACAGGGGACAGCCAGGCAATTCACGCTGTCTGACGCGTTCCCAACAT	1200		
QY	1201	GGACTCTTCAACACCACTGGGCGATGTCCCGCTGACGCCACCGGCTGTGTGGGCTGT	1260		
Db	1201	GGACTCTTCAACACCACTGGGCGATGTCCCGCTGACGCCACCGGCTGTGTGGGCTGT	1260		
QY	1261	GGTGTGTGGCAGGCACTGGGGCGGCGAGGAGAAAGCAGGCTTTCAGGAGCAGTCCGCG	1320		
Db	1261	GGTGTGTGGCAGGCACTGGGGCGGCGAGGAGAAAGCAGGCTTTCAGGAGCAGTCCGCG	1320		
QY	1321	GAGGAGTGCAACGAGGAGCCGGGCACTGCTGCTTCCCTGATGCTGACCCCAATTTGTC	1380		
Db	1321	GAGGAGTGCAACGAGGAGCCGGGCACTGCTGCTTCCCTGATGCTGACCCCAATTTGTC	1380		
QY	1381	TCCAGCCAGGCTTGGCAGAGCTGAGCACTGCAATGCAACAGTCTGGGTCAGATTGAT	1440		
Db	1381	TCCAGCCAGGCTTGGCAGAGCTGAGCACTGCAATGCAACAGTCTGGGTCAGATTGAT	1440		
QY	1441	ATCCGGGGCAGCTGCCCTGCCAAGCTGATGCCCGGCTATGGGCCCCCGGGATGCAGGC	1500		
Db	1441	ATCCGGGGCAGCTGCCCTGCCAAGCTGATGCCCGGCTATGGGCCCCCGGGATGCAGGC	1500		
QY	1501	CAGCAGAAAGGAAATCAACACAGAAAACGCCCCAACTTCCACAACTTCTGCAATGGCGAC	1560		
Db	1501	CAGCAGAAAGGAAATCAACACAGAAAACGCCCCAACTTCCACAACTTCTGCAATGGCGAC	1560		
QY	1561	ACCCAGAGACCAAGAGCATCAAGAGGAGACCCCGGATTCGGCTGAGACCCCGAGAG	1620		
Db	1561	ACCCAGAGACCAAGAGCATCAAGAGGAGACCCCGGATTCGGCTGAGACCCCGAGAG	1620		
QY	1621	GACCGTGTGGCGAGGGCCCTTGTCTTCTCTCTCTCTCTCTCTCTCTCTCTCTTACC	1680		
Db	1621	GACCGTGTGGCGAGGGCCCTTGTCTTCTCTCTCTCTCTCTCTCTCTCTCTCTTACC	1680		
QY	1681	GGGGTCAAACTCTGCTTTGGGCCATGAGCGAAATACATGGCCCTTCGCCCCCGTCACTCCG	1740		
Db	1681	GGGGTCAAACTCTGCTTTGGGCCATGAGCGAAATACATGGCCCTTCGCCCCCGTCACTCCG	1740		
QY	1741	GCCTGCCAGGTGATGACCGCATCACCAATCTCTGAGCAGCATTTATCGACAGGTGGTG	1800		
Db	1741	GCCTGCCAGGTGATGACCGCATCACCAATCTCTGAGCAGCATTTATCGACAGGTGGTG	1800		
QY	1801	GAAACGGAAGATCCAGGAGAAAGCCCTGGGGCCGGGGCTTCGAGCTGGCCCGGGTCTGGCG	1860		
Db	1801	GAAACGGAAGATCCAGGAGAAAGCCCTGGGGCCGGGGCTTCGAGCTGGCCCGGGTCTGGCG	1860		
QY	1861	AAGGGCTGGGCTGTCCCTCTCTCCAGTGGGCCCGGGCTGCTCCCGAGGGGCTTTG	1920		
Db	1861	AAGGGCTGGGCTGTCCCTCTCTCCAGTGGGCCCGGGCTGCTCCCGAGGGGCTTTG	1920		
QY	1921	CTGTGCTCAGAGGCCCGAGCTTGCCTCTGGCGTGGCTTCCACCTCTTCCAGGAGCAC	1980		
Db	1921	CTGTGCTCAGAGGCCCGAGCTTGCCTCTGGCGTGGCTTCCACCTCTTCCAGGAGCAC	1980		
QY	1981	TGAGCAGAGGCCAGCTGT	2040		
Db	1981	TGAGCAGAGGCCAGCTGT	2040		
QY	2041	TGGGGACAGAAAGCTCTTGGGGCACTTGGAGGCCAGGTGCAGGCGCTGAGCCCCCTCGGA	2100		



Db 2041 TGGGGGACAGAGCTCTTTGGGGCACTTGGAGCCAGGTGACGGGCTGAGCCCCCTCGGA 2100  
 Qy 2101 CCTCCAGCCAGCAGCCTGGGAGCACAACATTTCTGGAGGGCTTCTCTGGCCTGAG 2160  
 Db 2101 CCTCCAGCCAGCAGCCTGGGAGCACAACATTTCTGGAGGGCTTCTCTGGCCTGAG 2160  
 Qy 2161 CTTGCCCAAGTCAGAGGGCTCTGTCTCTCTGTGCAACGAGCTTTGGGGATGAG 2220  
 Db 2161 CTTGCCCAAGTCAGAGGGCTCTGTCTCTCTGTGCAACGAGCTTTGGGGATGAG 2220  
 Qy 2221 GACACGAGGGTGGAGAACTAGTCCAGTCTGCCACTTCCGAGTACTGCGCCCTC 2280  
 Db 2221 GACACGAGGGTGGAGAACTAGTCCAGTCTGCCACTTCCGAGTACTGCGCCCTC 2280  
 Qy 2281 CATGAAACTCAACTGGCTTCTTACCTCCACCGGGCTTGGCCCTGCTCCACTGGAG 2340  
 Db 2281 CATGAAACTCAACTGGCTTCTTACCTCCACCGGGCTTGGCCCTGCTCCACTGGAG 2340  
 Qy 2341 CCCAGCTCTGGGAGCCATGTGTGAGCCGCGACCGGGACACCTGGGACCAAGAAC 2400  
 Db 2341 CCCAGCTCTGGGAGCCATGTGTGAGCCGCGACCGGGACACCTGGGACCAAGAAC 2400  
 Qy 2401 CTCTGTGGAGGTGGCGGACTGTGTGAGCATCTGTGTGATGCGACACACCACTGCCT 2460  
 Db 2401 CTCTGTGGAGGTGGCGGACTGTGTGAGCATCTGTGTGATGCGACACACCACTGCCT 2460  
 Qy 2461 GCCTGGCACCGGGACAGAAAGACTTCTTTTTCAGGCTTGGAGCGGGGGCTGTGTCT 2520  
 Db 2461 GCCTGGCACCGGGACAGAAAGACTTCTTTTTCAGGCTTGGAGCGGGGGCTGTGTCT 2520  
 Qy 2521 CCGGAGCAGCAGGTGAGCAGTGTGTGAGCGTGTTCGGGACAGGACGCCAGCGATC 2580  
 Db 2521 CCGGAGCAGCAGGTGAGCAGTGTGTGAGCGTGTTCGGGACAGGACGCCAGCGATC 2580  
 Qy 2581 CGCGCTTCTTCCAGATGCTGCGCGGCGGGGAGCGGCTGTGAGCTGGCGGCCCA 2640  
 Db 2581 CGCGCTTCTTCCAGATGCTGCGCGGCGGGGAGCGGCTGTGAGCTGGCGGCCCA 2640  
 Qy 2641 GCGAGCTCTACCTGATGAGGCTGCGGGGCGGCTTGGAGGAGTGGGGCTGAGC 2700  
 Db 2641 GCGAGCTCTACCTGATGAGGCTGCGGGGCGGCTTGGAGGAGTGGGGCTGAGC 2700  
 Qy 2701 TGCTGGACCTTCTCAGGCCCCCGGAGAGCGCTGTGTGCTGTGAGGGCTTCCAC 2760  
 Db 2701 TGCTGGACCTTCTCAGGCCCCCGGAGAGCGCTGTGTGCTGTGAGGGCTTCCAC 2760  
 Qy 2761 CAGGTGAGGGGCTGTGTGAGCAGTCACTGAGCTTCTTCTTCCCTGAGACC 2820  
 Db 2761 CAGGTGAGGGGCTGTGTGAGCAGTCACTGAGCTTCTTCTTCCCTGAGACC 2820  
 Qy 2821 TCTGCCCTCTCTGCTCAGCTTCCACAGGAGCCAGCTTCCCTTCTGCTGACCTG 2880  
 Db 2821 TCTGCCCTCTCTGCTCAGCTTCCACAGGAGCCAGCTTCCCTTCTGCTGACCTG 2880  
 Qy 2881 CTTATGCCAGATGAGCTGGCTGTGTTCAGAGTGAAGTGGCGTGGGACATTA 2940  
 Db 2881 CTTATGCCAGATGAGCTGGCTGTGTTCAGAGTGAAGTGGCGTGGGACATTA 2940  
 Qy 2941 CAGGAGGCCAATAGAGGATCTAGGTGTCTGGGATCGGGTGGGACAGGTAGACAG 3000  
 Db 2941 CAGGAGGCCAATAGAGGATCTAGGTGTCTGGGATCGGGTGGGACAGGTAGACAG 3000  
 Qy 3001 GTGCTCAGCCAGGACCACTTTCAGAGGGATGGCGCTAGGGGACTTGGGATTTCTGG 3060  
 Db 3001 GTGCTCAGCCAGGACCACTTTCAGAGGGATGGCGCTAGGGGACTTGGGATTTCTGG 3060  
 Qy 3061 TCAACCCACAAAGCAGTCTGTGGGACAAAGCAGGAGCACTGTGTTCCCTCCCTTTAG 3120  
 Db 3061 TCAACCCACAAAGCAGTCTGTGGGACAAAGCAGGAGCACTGTGTTCCCTCCCTTTAG 3120  
 Qy 3121 CCAACACACAGTGCACCAAGCTCACACCTGTCTTCTCAGGTGGCATCTCCCCAC 3180

Db 3121 CCAACACACAGTGCACCAAGCTCACACCTGTCTTCTCAGGTGGCATCTCCCCAC 3180  
 Qy 3181 CCTGTGCCCTTTTATGTACAGG 3202  
 Db 3181 CCTGTGCCCTTTTATGTACAGG 3202

RESULT 2  
 AF039196 5709 bp mRNA linear PRI 15-APR-2002  
 LOCUS Homo sapiens putative single zinc finger transcription factor  
 DEFINITION protein (hairless) mRNA, complete cds.  
 ACCESSION AF039196  
 VERSION AF039196.3 GI:20149786  
 KEYWORDS  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 REFERENCE 1 (bases 1 to 5709)  
 AUTHORS Ahmad, W., ul Haque, M.F., Brancolini, V., Tsou, H.C., ul Haque, S.,  
 Lam, H., Aita, V.M., Owen, J., deBlaquiere, M., Frank, J.,  
 Cserhalmi-Friedman, P.B., Leask, A., McGrath, J.A., Peacocke, M.,  
 Ahmad, M., Ott, J. and Christiano, A.M.  
 TITLE Alopecia universalis associated with a mutation in the human  
 hairless gene  
 JOURNAL Science 279 (5351), 720-724 (1998)  
 MEDLINE 98111413  
 PUBMED 9445480  
 REFERENCE 2 (bases 1 to 5709)  
 AUTHORS Ahmad, W., Zlotogorski, A., Panteleyev, A.A., Lam, H., Ahmad, M., ul  
 Haque, M.F., Abdallah, H.M., Dragan, L. and Christiano, A.M.  
 TITLE Genomic organization of the human hairless gene (HR) and  
 identification of a mutation underlying congenital atrichia in an  
 Arab Palestinian family  
 JOURNAL Genomics 56 (2), 141-148 (1999)  
 MEDLINE 99162400  
 PUBMED 10051399  
 REFERENCE 3 (bases 1 to 5709)  
 AUTHORS Christiano, A.M.  
 TITLE Direct Submission  
 JOURNAL Submitted (18-DEC-1997) Dermatology, Columbia University, 630 West  
 168th Street VC-1526, New York, NY 10032, USA  
 REFERENCE 4 (bases 1 to 5709)  
 AUTHORS Christiano, A.M.  
 TITLE Direct Submission  
 JOURNAL Submitted (18-AUG-1998) Dermatology, Columbia University, 630 West  
 168th Street VC-1526, New York, NY 10032, USA  
 REFERENCE 5 (bases 1 to 5709)  
 AUTHORS Christiano, A.M.  
 TITLE Direct Submission  
 JOURNAL Submitted (11-JUN-1999) Dermatology, Columbia University, 630 West  
 168th Street VC-1526, New York, NY 10032, USA  
 REFERENCE 6 (bases 1 to 5709)  
 AUTHORS Christiano, A.M.  
 TITLE Direct Submission  
 JOURNAL Submitted (15-APR-2002) Dermatology, Columbia University, 630 West  
 168th Street VC-1526, New York, NY 10032, USA  
 REFERENCE 7 (bases 1 to 5709)  
 AUTHORS Christiano, A.M.  
 TITLE Direct Submission  
 JOURNAL Submitted (15-APR-2002) Dermatology, Columbia University, 630 West  
 168th Street VC-1526, New York, NY 10032, USA  
 COMMENT On Apr 15, 2002 this sequence version replaced gi:5042382.  
 FEATURES  
 Location/Qualifiers  
 1..5709  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /chromosome="8"  
 /map="8p12"  
 1..5709  
 /gene="hairless"  
 1485..5054  
 /gene="hairless"

source  
 gene  
 CDS

```
/note="restricted expression in the brain and skin"
/codon_start=1
/product="putative single zinc finger transcription factor
protein"
/protein_id="AAC32258.3"
/db_xref="GI:2014978"
/translation="MESTPFLKGTPTWEKTAPENGIVROEPGSPRRDLHGFLCLG
EPAPFRWGLVSTDSWLPFGPPQPKDMLPLVEGEGPQNGERKVNWLGSKELRWKA
MLTPELAFCGPACPRCPDMLTPHPWVYSGQPKVPSAFSLGSKGVYKQPSLRLAK
EPLAAAEPLGLNSGHLORAGAEARPSLHORDEMGAGRQNPCLFLQOPTVPM
TSWAPCPGLVHITLGNWAGPGLTQUGPPATPCPSPPEPVTQRCGCCSYPTK
GSDPGCKQEGDEGLSGASEPSEVNKASGPAPCPSPSHHTKLKTLWLRHSEQPE
CRCPVEVERPVARLALKEAGSEVQAGMSPAPKPPDPFPPTAEQAGAGWQEV
DTISGNKVDSDGDEKQPDQGLQDIPCLALPAKLAQCSQCAOAAEGEG
GHACHSOQVRRSLPGLEQDEDTATNSSEEGPGSPDSRLSTGLAKHLLSLGLDL
CLLRREAREALWAQSRGQGPANTEDESPGIPRCRCHHGLFNTHWRPCRSRLCLVA
CGRVAGTRAREKAFQESAECTQAGHAACSLMLTFVSSQALAEKSTAMHQVW
KFDLRGHCPOADARWAPDAGQOKSTQKTPPTQPCNGDTHRTKSIKEETPDSA
ETPAEDRAGRGLPCPSLCHELLASTAVKLCLGHERIHMAFAPVTPALPSDDRTINLD
SIIAQVVERKIQERALGPRAGRLKGLGLPLSPVRPRLPPPGALLWLQEPQPCPR
RGPHLFOEHRKQOPVLVSGTORTLQNLWGLTEALGALGQVQALSPLGPQPSLGS
TTWEGFSEWLEPKSDGSLVLLHRLALGDEDTSRVENLAASLPLPHYCALHKLNL
SYLPPGLALPLEPQLWAAYGVSPHRLGLTKNLCEVADLVLSLHADTLPFAWHR
KQPLSLGDEGLMSPGSQVSTVWHVFRQAQRIIRFLQWVCAGAGALEPAGPASC
YLDAGLRRLREEWGVSCTLLQAPGEAVLVPAGAPHQVGLVSTVSTQHFLEPETS
ALSAQLCHQPSLPPDCHLLYAQMDWAVFOAVKVAVGTLQFAK"
```

## ORIGIN

```
Query Match 99.5%; Score 3185; DB 9; Length 5709;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 3188; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 1 TTTTACTACAGGATCCGAGCATCCACAGTTGCAAGAGGACCTTGCGAGCTGCGGAA 60
Db 2100 TTTTACTACAGGATCCGAGCATCCACAGTTGCAAGAGGACCTTGCGAGCTGCGGAA 2159

Qy 61 CTTGGTGTGTTTGTCTTAACTCTGGTGGGACCTCGAGAGAGCGGGGAGCGGAAACGC 120
Db 2160 CTTGGTGTGTTTGTCTTAACTCTGGTGGGACCTCGAGAGAGCGGGGAGCGGAAACGC 2219

Qy 121 CTTTCACTGCACAGAGGATGGAGATGGAGCTGGCGGCGAGAGATCTTGTGCGG 180
Db 2220 CTTTCACTGCACAGAGGATGGAGATGGAGCTGGCGGCGAGAGATCTTGTGCGG 2279

Qy 181 CTCTTCTCGGGGAGCAGACACTGTGCGCTTGGACCTCTTGGCGCGCTTGTCCCGCAGGC 240
Db 2280 CTCTTCTCGGGGAGCAGACACTGTGCGCTTGGACCTCTTGGCGCGCTTGTCCCGCAGGC 2339

Qy 241 CTTGTTTCACTCTTGGCAACGCTTGGGCTGGGCGAGCGATGGGAACTTGGGTACCA 300
Db 2340 CTTGTTTCACTCTTGGCAACGCTTGGGCTGGGCGAGCGATGGGAACTTGGGTACCA 2399

Qy 301 CTGGGCGCACAGCAACCAAGTGCCTCTCTCTGAGCGGCTGTCAACCGCGGGG 360
Db 2400 CTGGGCGCACAGCAACCAAGTGCCTCTCTCTGAGCGGCTGTCAACCGCGGGG 2459

Qy 361 TGCTGTTTCATCTTACCACCACTTAAAGTGGGATCTTGGCCCTTGTGGGAAGTGCAG 420
Db 2460 TGCTGTTTCATCTTACCACCACTTAAAGTGGGATCTTGGCCCTTGTGGGAAGTGCAG 2519

Qy 421 GAGGCGCTGAGGGGGGTGCCAGTGGAGCGCAGGAAACCCAGCGAGGAAGTGAACAAGGCC 480
Db 2520 GAGGCGCTGAGGGGGGTGCCAGTGGAGCGCAGGAAACCCAGCGAGGAAGTGAACAAGGCC 2579

Qy 481 TCTGGCCCCAGGGGCTGTCCCCCAGCCACACCAAGCTGAAGAGAGATGCTCACA 540
Db 2580 TCTGGCCCCAGGGGCTGTCCCCCAGCCACACCAAGCTGAAGAGAGATGCTCACA 2639

Qy 541 CGGCACCTCGAGCAGGTTGAATGTCCACGCGGCTGCGCTGAGGTCGAGGAGAGCGCGTT 600
Db 2640 CGGCACCTCGAGCAGGTTGAATGTCCACGCGGCTGCGCTGAGGTCGAGGAGAGCGCGTT 2699
```

```
Qy 601 GCTCGCTCCGGGCCCTCAAAAGGCGACGAGCCCGAGGTCACAGGAGCAATGGGCAGT 660
Db 2700 GCTCGCTCCGGGCCCTCAAAAGGCGACGAGCCCGAGGTCACAGGAGCAATGGGCAGT 2759

Qy 661 CCAGCCCCAAGCGGCGACCGGACCTTTTCCAGGCACTGCAGAACAGGGGGCTGGGGGT 720
Db 2760 CCAGCCCCAAGCGGCGACCGGACCTTTTCCAGGCACTGCAGAACAGGGGGCTGGGGGT 2819

Qy 721 TGGCAGAGAGTTCCGGGACACATCATAGGNAACAGAGATGTGACTCGGAGACACATCAT 780
Db 2820 TGGCAGAGAGTTCCGGGACACATCATAGGNAACAGAGATGTGACTCGGAGACACATCAT 2879

Qy 781 GAGCAGAAAGACCCCAAGATGGCAGGCCAGTCTTCAGGACCCCGGAGCTTCAGACATA 840
Db 2880 GAGCAGAAAGACCCCAAGATGGCAGGCCAGTCTTCAGGACCCCGGAGCTTCAGACATA 2939

Qy 841 CCATGCTGCTTCTCCCTGTGAAAACCTGGCTCAATGTCGCAAAAGTTGTGCCAGGCACTGGA 900
Db 2940 CCATGCTGCTTCTCCCTGTGAAAACCTGGCTCAATGTCGCAAAAGTTGTGCCAGGCACTGGA 2999

Qy 901 GAGGAGGAGGCGACGCTGCACTCTCAGCAAGTGGGAGATCGCTCTGGGAGGGAG 960
Db 3000 GAGGAGGAGGCGACGCTGCACTCTCAGCAAGTGGGAGATCGCTCTGGGAGGGAG 3059

Qy 961 CTCGAGCAGGAGGAGACACAGCCACCACTCTCAGGAGGAGGAGGAGGAGGAGGAGGAG 1020
Db 3060 CTCGAGCAGGAGGAGGAGACACAGCCACCACTCTCAGGAGGAGGAGGAGGAGGAGGAGGAG 3119

Qy 1021 CCTGACAGCGGCTCAGCAGAGGCTGCCNAGCACCTGCTCAGTGGTTTGGGGAGCCGA 1080
Db 3120 CCTGACAGCGGCTCAGCAGAGGCTGCCNAGCACCTGCTCAGTGGTTTGGGGAGCCGA 3179

Qy 1081 CTGTGCGGCTGCTGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1140
Db 3180 CTGTGCGGCTGCTGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3239

Qy 1141 GGGCAGCCGCTGACAGAGGAGCAGCCAGGCACTTCCAGCTGCTCAGCGCTTGGCCACCAT 1200
Db 3240 GGGCAGCCGCTGACAGAGGAGCAGCCAGGCACTTCCAGCTGCTCAGCGCTTGGCCACCAT 3299

Qy 1201 GGACTCTTAAACCCACTGGCGATGTCCCGCTGAGGAGCAGCGGCTGTGTGGGCTGTGT 1260
Db 3300 GGACTCTTAAACCCACTGGCGATGTCCCGCTGAGGAGCAGCGGCTGTGTGTGGGCTGTGT 3359

Qy 1261 GGTGCTGTGAGGAGCAGTGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1320
Db 3360 GGTGCTGTGAGGAGCAGTGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3419

Qy 1321 GAGGAGTGCAAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1380
Db 3420 GAGGAGTGCAAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3479

Qy 1381 TCCAGCCAGGCTTGGGAGAGCTGAGCAGTGCATGCAATGCAACAGAGCTCTGGTCAAGTTGAT 1440
Db 3480 TCCAGCCAGGCTTGGGAGAGCTGAGCAGTGCATGCAATGCAACAGAGCTCTGGTCAAGTTGAT 3539

Qy 1441 ATCCGGGGGCACTGCCCTGCAAGCTGATGCCGGGTATGGGCGCCCGGGAGTGCAGGC 1500
Db 3540 ATCCGGGGGCACTGCCCTGCAAGCTGATGCCGGGTATGGGCGCCCGGGAGTGCAGGC 3599

Qy 1501 CAGCAGAGGAATCAACACAGAGAAACGCCGCCCACTCCCAAACTTCTCTGCAATGGCGAC 1560
Db 3600 CAGCAGAGGAATCAACACAGAGAAACGCCGCCCACTCCCAAACTTCTCTGCAATGGCGAC 3659

Qy 1561 ACCCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1620
Db 3660 ACCCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3719

Qy 1621 GAGCGCTGCTGGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1680
Db 3720 GAGCGCTGCTGGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3779

Qy 1681 GCGGTCAAACTGTGTGGGCGCATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1740
```



5'UTR	/gene="HR" 1..40		Db	1076	GAGGGCTGTGAGGGGGTGTGCTGAGGCGAGCGAA
CDS	/gene="HR" 41..3610		QY	481	TCCTGGCCCCAGGGCTGTCCCCCAGCCACCA
	/gene="HR"		Db	1136	TCCTGGCCCCAGGGCTGTCCCCCAGCCACCA
	/note="putative single zinc finger transcription factor protein, responsible for autosomal recessive congenital alopecia"		QY	541	CGGCACCTCGGAGCAGTTTGAATGTCCACGCGGTGCTGCTGAGGTTCGAGGAGAGCGCGGTT
	/codon_start=1		Db	1196	CGGCACCTCGGAGCAGTTTGAATGTCCACGCGGTGCTGCTGAGGTTCGAGGAGAGCGCGGTT
	/product="hairless protein"		QY	601	GCTCGGCTCGGGCTCTAAAGGGGAGGAGCGAGCCCGAGGTCCAGGAGCAATGGGAGT
	/protein_id="CAB86602.1"		Db	1256	GCTCGGCTCGGGCTCTAAAGGGGAGGAGCGAGCCCGAGGTCCAGGAGCAATGGGAGT
	/db_xref="GI:7529560"		QY	661	CCAGCCCCAAGCGGCGACCGGACCTTTTCCAGGGCACTGCAGAA
	/db_xref="GOA:O43593"		Db	1316	CCAGCCCCAAGCGGCGACCGGACCTTTTCCAGGGCACTGCAGAA
	/translation="MESTPSFLKGTPTWEKTAPENGIVROBSPGPRDGLHGPCLIGEPAPFWRGVLSLTPDPQPGKMDLVEGEGPONGERKVNVLKSGEGRWKEA		QY	721	TGGCAGAGGTTTCGGGACACATCGATAGGAA
	MLTHPLAFVCGPACPPRCPPGSGGHLKSDPVAFPMHCPLELLEKILERRFPWP		Db	1376	TGGCAGAGGTTTCGGGACACATCGATAGGAA
	TCLPYLVSLGPPHPCDMLTTPHWYSGQPKVPSAFLGSKGFYKDPSPRLAK		QY	781	GAGCAGAAAGGACCCCAAGATGTCAGAGCCCGGATCTTCAGGACATA
	EPLAAAPGLFCLNSGHLRAGAEAPSLHQRDEMGARQNPCLFLGQDPDTPW		Db	1436	GAGCAGAAAGGACCCCAAGATGTCAGAGCCCGGATCTTCAGGACATA
	TSWAPCPGLVHTLGNVAGPNDNLGYLGPATPRCPSPFPVPTGCGCCSYPTTK		QY	841	CMATGCTCTTCTCCCTCAAACTCGCTCAATGCGGAGATCGCTTCGGAGGGGAG
	GGGPGCGQEGLEGGASGAPSEEVNKASGPACPPSHHTKLTMTLRHSEBPQ		Db	1496	CMATGCTGCTCTCCCTGCAAACTGCGCTCAATGCGGAGATCGCTTCGGAGGGGAG
	CPRCPVEEPRVARALKRAGSPVQGVGAMSPAPKRPDPFGTAEQAGGQWVR		QY	901	GAGGAGAGGGGACGCTGCCACTCTCAGCAAGTTCGCGAGATCGCTTCGGAGGGGAG
	DTISGNVDGSDGDEQKQDQASLDPLQDIPCLAPLAKLAQSCAQAAGBGR		Db	1556	GAGGAGAGGGGACGCTGCCACTCTCAGCAAGTTCGCGAGATCGCTTCGGAGGGGAG
	GHACSOQVRSPLGSELOEBEDATNSESSEPGSDPSDELSTGLAKHLLSLGRV		QY	961	CTGCAGCAGAGGAGACACAGCCACCAACTCGACTTGTAGGAAAGGCCAGCGTCCGC
	CRLRRRREALNAWRGQGPATVDSFGPIRCCSRCHGLFNTHRCPCRSRLCVA		Db	1616	CTGCAGCAGAGGAGACACAGCCACCAACTCGACTTGTAGGAAAGGCCAGCGTCCGC
	CGRVATGRARAGFQEGSAECTQEGAAHACSLMTQFVSQAALASHTAMHVW		QY	1021	CTGCAGCAGCGCTCAGACAGCGCTCGCCAGACCTGCTCAGTGTGTGGGGACCGA
	KFDIRGTCOARAWAPADQKQESTQTPPTPQPSNCGDTHRTKSIKEBTPQ		Db	1676	CTGCAGCAGCGCTCAGACAGCGCTCGCCAGACCTGCTCAGTGTGTGGGGACCGA
	ETPAEDRAGRPPLPCLSCHELLASTAVKLCLGHERIHMAFPVTPALPSDRITNLD		QY	1081	CTGTGCGGCTCTGTCGGAGGAGCGGAGGCGCTTGGTTCAGTGTGTGGGGACCGA
	SIIAUVVERIKQKALGPRAGRLKGLHSPVRPRLPPGALLMLQEPQCFR		Db	1736	CTGTGCGGCTCTGTCGGAGGAGCGGAGGCGCTTGGTTCAGTGTGTGGGGACCGA
	RGHFLFOHWRQOPVLVSGIORTLQNLGTEALGALGQVQALSPLPQPSLGS		QY	1141	GGGCGAGCGCTGACAGGGGACAGCCACAGCAATTCACGCTGCTGAGCGGTTGCCACCAT
	SYLPPGLALRPLBPQLMAAYGVSPHRLGHTKNLCEVADLVSLVHADTPLPAMHRA		Db	1796	GGGCGAGCGCTGACAGGGACAGCCACAGCAATTCACGCTGCTGAGCGGTTGCCACCAT
	QKDFSLDGLGWSQVSTVWVFRQAQRIIRFLQWCPAGAGALEPAGPAGSC		QY	1201	GGACTCTTCAACACCCACTGGCGATGTCCCCTGTCAGCCACCGCTGTGTGCGCTGT
	YLDAGLRIRLEEWGVSCTWLLQAPGEAVLVPAGAPHOVLGVSTVSTVTHPLSPETS		Db	1856	GGACTCTTCAACACCCACTGGCGATGTCCCCTGTCAGCCACCGCTGTGTGCGCTGT
	ALSAQLCHQGPSLPDCHLLYAQMDMAVFQAVKAVGTLQEA"		QY	1261	GCTGCTGTGGCAGGACCTGGGCGGCGCAGGGAGAAAGCAGGCTTTTCAGGAGCAGTCCGG
	3611..>3719		Db	1916	GCTGCTGTGGCAGGACCTGGGCGGCGCAGGGAGAAAGCAGGCTTTTCAGGAGCAGTCCGG
	/gene="HR"		QY	1321	GAGGAGTGACACGAGAGCGGGGACGCTGCTGCTTCCCTGATGCTGACCCAGTTTGTGTC
			Db	1976	GAGGAGTGACACGAGAGCGGGGACGCTGCTGCTTCCCTGATGCTGACCCAGTTTGTGTC
			QY	1381	TCGACCGCAGGCTTTTGGCAGAGCTGAGCACTGCAATGCAACAGGTCTGGGTCAAGTTTGAAT
			Db	2036	TCGACCGCAGGCTTTTGGCAGAGCTGAGCACTGCAATGCAACAGGTCTGGGTCAAGTTTGAAT
			QY	1441	ATCCCGGGGCACTGCCCCCTTGCCAAAGCTGATGCCCGGGTATGGGCCCCCGGGGATGCAGGC
			Db	2096	ATCCCGGGGCACTGCCCCCTTGCCAAAGCTGATGCCCGGGTATGGGCCCCCGGGGATGCAGGC
			QY	1501	CAGCAGAGGATCAACACAGAAAGCCCCCACTCCACAACTTCTCTCAATGGGAC
			Db	2156	CAGCAGAGGATCAACACAGAAAGCCCCCACTCCACAACTTCTCTCAATGGGAC
3'UTR			QY	421	GAGGGCTGTGAGGGGGTGTGCTGAGCAGCGAACCAGCGAGGAAGTGTACAGGCC
ORIGIN			Db		
Query Match	95.0%; Score 3042.4; DB 9; Length 3719;				
Best Local Similarity	99.8%; Pred. No. 0;				
Matches 3057; Conservative	0; Mismatches 6; Indels 1; Gaps 1;				
QY	1	TTTTACTAAGAGTCCAGCATCCAGTTTGGCAAGGAGCCCTTGGCAGTGGGAA	60		
Db	656	TTTTACTAAGAGTCCAGCATCCAGTTTGGCAAGGAGCCCTTGGCAGTGGGAA	715		
QY	61	CCTGGGTTGTTGGCTTAACTCTGTGGGCACCTGCAGAGCCGGGAGGCGCAACGC	120		
Db	716	CCTGGGTTGTTGGCTTAACTCTGTGGGCACCTGCAGAGCCGGGAGGCGCAACGC	775		
QY	121	CCTTCACTCACCAGAGGATGAGAGATGGGAGCTGGCGCGCAGCAGATCTCTGCCCG	180		
Db	776	CCTTCACTCACCAGAGGATGAGAGATGGGAGCTGGCGCGCAGCAGATCTCTGCCCG	835		
QY	181	CTCTTCTGGGGCAGCCAGACACTGTGCCCTGACCTCTTGGCCCGCTTGTCCCCCAGGC	240		
Db	836	CTCTTCTGGGGCAGCCAGACACTGTGCCCTGACCTCTTGGCCCGCTTGTCCCCCAGGC	895		
QY	241	CTTGTTTCATCTCTTGCACAGCTCTGGCTGGGCGCAGGCGATGGAACTTGGGTACAG	300		
Db	896	CTTGTTTCATCTCTTGCACAGCTCTGGCTGGGCGCAGGCGATGGAACTTGGGTACAG	955		
QY	301	CTGGGGCCACAGCAACACCAAGGTGCCCTCTCTTGGCCGCTGT			

1561	QY	ACCCAGGACCAAGAGCATCAAGAGGAGACCCCGATTCCGCTGAGACCCCGAGAG	1620
2216	Db	ACCCAGAGACCAAGAGCATCAAAGAGGAGACCCCGATTCCGCTGAGACCCCGAGAG	2275
1621	QY	GACCGTGCTGGCCGAGGGGCCCTTGCCCTTGCTCTCTCTGCGAACTGCTGGCTTCTACC	1680
2276	Db	GACCGTGCTGGCCGAGGGGCCCTTGCCCTTGCTCTCTCTGCGAACTGCTGGCTTCTACC	2335
1681	QY	GGGCTCAAACTCTGCTTGGGCCATGAGCGAATAACAATGGCCTTCGCCCCGCTCACTCCG	1740
2336	Db	GGGCTCAAACTCTGCTTGGGCCATGAGCGAATAACAATGGCCTTCGCCCCGCTCACTCCG	2395
1741	QY	GCCTCCCACTGATGACCGCATCACCAATCCTTGACAGCAATATCGCACAGGTGGTG	1800
2396	Db	GCCTGCCCACTGATGACCGCATCACCAATCCTTGACAGCAATATCGCACAGGTGGTG	2455
1801	QY	GAAACGGAAGATCCAGGAGAAAGCCCTTGGGCCCGGGGCTTCGAGCTGGCCCGGCTTCGCGC	1860
2456	Db	GAAACGGAAGATCCAGGAGAAAGCCCTTGGGCCCGGGGCTTCGAGCTGGCCCGGCTTCGCGC	2515
1861	QY	AAGGGCCTGGGCCCTGGCCCTCTCTCAGTGGCGCCCGGCTGCTCTCCCGAGGGGCTTTG	1920
2516	Db	AAGGGCCTGGGCCCTGGCCCTCTCTCAGTGGCGCCCGGCTGCTCTCCCGAGGGGCTTTG	2575
1921	QY	CTGTGGCTGCAGAGGCCCGAGCCCTTCGCTCGGCGTGGCTTCCACCTCTTCACAGGAGCAC	1980
2576	Db	CTGTGGCTGCAGAGGCCCGAGCCCTTCGCTCGGCGTGGCTTCCACCTCTTCACAGGAGCAC	2635
1981	QY	TGAGGSCAGGGCAGCCTGTGTTGGTGTCAGGGATCCAAAGACATTTGAGGGCAACCTTG	2040
2636	Db	TGAGGSCAGGGCAGCCTGTGTTGGTGTCAGGGATCCAAAGACATTTGAGGGCAACCTTG	2695
2041	QY	TGGGGACACAAAGCTCTTGGGGCACTTGGAGGCAAGTGCAGGCGCTGAGCCCTCGGA	2100
2696	Db	TGGGGACACAAAGCTCTTGGGGCACTTGGAGGCAAGTGCAGGCGCTGAGCCCTCGGA	2755
2101	QY	CTTCCCAGCCACAGAGCCTGGGACACAAATCTTGGGAGGGCTTCTCTGGCCCTGAG	2160
2756	Db	CTTCCCAGCCACAGAGCCTGGGACACAAATCTTGGGAGGGCTTCTCTGGCCCTGAG	2815
2161	QY	CTTGGCCCAAGTCAGACGAGGGCTGTCTCTCTCTGTCACCGAGCTTTGGGGGATGAG	2220
2816	Db	CTTGGCCCAAGTCAGACGAGGGCTGTCTCTCTCTGTCACCGAGCTTTGGGGGATGAG	2875
2221	QY	GACACGACGAGGTGAGAACTAGCTGCCGCTGCGCACTCCGAGTACTTGCGCCCTC	2280
2876	Db	GACACGACGAGGTGAGAACTAGCTGCCGCTGCGCACTCCGAGTACTTGCGCCCTC	2935
2281	QY	CATGHAATACTCAACTGGCTTCTCTACCTCCACCGGGGCTTGCCCTGCGTCCACTGGAG	2340
2936	Db	CATGHAATACTCAACTGGCTTCTCTACCTCCACCGGGGCTTGCCCTGCGTCCACTGGAG	2995
2341	QY	CCGACGCTTGGGCAAGCTATGGTGAGCCGCAACCGGGGACACTTGGGGACCAAGAATC	2400
2996	Db	CCGACGCTTGGGCAAGCTATGGTGAGCCGCAACCGGGGACACTTGGGGACCAAGAATC	3055
2401	QY	CTCTGTGTGAGGTGCGACCTGGTCAGCATCTGGTGCATGCGCACACCACTGGCT	2460
3056	Db	CTCTGTGTGAGGTGCGGACCTGGTCAGCATCTGGTGCATGCGCACACCACTGGCT	3115
2461	QY	GCCTGGCACCGGSCACAGAAAGACTTCTTTTTCAGGCTTCGAGCGGGAGGGCTCTGGTCT	2520
3116	Db	GCCTGGCACCGGSCACAGAAAGACTTCTTTTTCAGGCTTCGAGCGGGAGGGGCTCTGGTCT	3175
2521	QY	CCGGGACGCAAGTCAAGACTGTGTGGCACTGTTCGGGCAACAGACGCCACGCGCATC	2580
3176	Db	CCGGGACGCAAGTCAAGACTGTGTGGCACTGTTCGGGCAACAGACGCCACGCGCATC	3235
2581	QY	CGCCGCTTTCTCCAGATGTGTCCCGCGCGGGGACGGCCCTTGAGCCTGGCGCCCA	2640
3236	Db	CGCCGCTTTCTCCAGATGTGTGCCCGCGCGGGGACGGCCCTTGAGCCTTCGCGCCA	3295

Qy	2641	GGCAGCTGCTACTGATGACAGGGCTGCGCGCGCCCTGCGGAGAGTGGGGCGTGAGC	2700
Db	3296	GGCAGCTGCTACTGATGACAGGGCTGCGCGCGCCCTGCGGAGAGTGGGGCGTGAGC	3355
Qy	2701	TGCTGGACCCCTGCTCAGAGCCCCGGAGAGGCGTGCTGTGCTGCTGCAGAGGGTCCCCAC	2760
Db	3356	TGCTGGACCCCTGCTCAGAGCCCCGGAGAGGCGTGCTGTGCTGCTGCAGAGGGTCCCCAC	3415
Qy	2761	CAGGTGAGGGCTGTGTAGACACAGTCAGCGTCACTCAGCACTTCCTCTCCCTCGAGACC	2820
Db	3416	CAGGTGAGGGCTGTGTAGACACAGTCAGCGTCACTCAGCACTTCCTCTCCCTCGAGACC	3475
Qy	2821	TCTGSCCTCTCTGCTCAGCTCTGCCACCAAGGACCCAGCCCTCCCCCTGACTGCCACCTG	2880
Db	3476	TCTGSCCTCTCTGCTCAGCTCTGCCACCAAGGACCCAGCCCTCCCCCTGACTGCCACCTG	3535
Qy	2881	CTTTATGCCCAGATGGACTGGGCTGTGTTCCAAAGCAGTGAAGGTGGCCGTGGGGACATT	2940
Db	3536	CTTTATGCCCAGATGGACTGGGCTGTGTTCCAAAGCAGTGAAGGTGGCCGTGGGGACATT	3595
Qy	2941	CAGAGGCCAAATAGAGGGATGTAGTGTCTGGGATCGGGGTGGGACAGGTAGACCAG	3000
Db	3596	CAGAGGCCAAATAGAGGGATGTAGTGTCTGGGATCGGGGTGGGACAGGTAGACCAG	3655
Qy	3001	GTGCTC-AGCCCAGGCACAACTTCAGCAGGGGATGCGCTAGCGGACTTCGGGATTTCTG	3059
Db	3656	GTGCTCAAGCCCAAGCACACTTCAGCAGGGGATGCGCTAGCGGACTTCGGGATTTCTG	3715
Qy	3060	GTCA 3063	
Db	3716	GTCA 3719	

RESULT 4  
AF361864  
LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL

3564 bp mRNA linear  
Macaca mulatta hairless mRNA, complete cds.  
AF361864.1 GI:18028978  
Macaca mulatta (rhesus monkey)  
Macaca mulatta  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;  
Cercopithecinae; Macaca.  
1 (bases 1 to 3564)  
Ahmad, W., Ratterre, M.S., Panteleyev, A.A., Aita, V.M., Sundberg, J.P.  
and Christiano, A.M.  
Papular Atrichia Resulting from Mutations in the Rhesus Monkey  
(Macaca mulatta) hairless Gene  
Unpublished  
2 (bases 1 to 3564)  
Ahmad, W., Ratterre, M.S., Panteleyev, A.A., Aita, V.M., Sundberg, J.P.  
and Christiano, A.M.  
Direct Submission  
Submitted (16-MAR-2001) Department of Dermatology, Columbia  
University, 630 W 168th St, New York, NY 10032, USA

CDS

GLGPCCKQCEGLDGVASGASEPSEBVKASGPRACPFGHHTKLKKTWLTRHSEQPECP  
RCPFAEERPAQLRALKRAGSPEIQGAVGGPAPKRPDPFPGTAEQAGAGWQEVDRTP  
STGNKEADSGVHDDQRPDRGOASLQDGLQRIPLCALPAKLAQCCQAQAAGEGGGP  
AGHPQVRRSPGLGELQEEEDTAANSSEEGPGSDGRLSTGLSKHLLSSLDRLCR  
LLRREKALAMARSGQGPVAVDTGPIPRCCSRCHGLFNTHWRCPRSRHLRCVACG  
RVAGARAFKESRESEETCTQAGHAACSLATLTFVSSQALAEELSAMEQVWKF  
DIRGHCPOADARVAPGDAGQKESTPTPTPOPCNGDTHRTKSIKEETPDSJET  
PAEDRAGQAPLPCPSLCELLASTAVKLCLGHERIHMAFAPVTPPLSDRIINLIDS  
TAOVVERIKOEKALGPGURAGPLRKLGLPLSPVRPLPPGALLMLDEPRPRQRG  
FHLFBHWRQGPVLVSQRTLOLGNLNGTEALGALGQVQALSPLGPQPTSLGSTA  
FWEGSWPELRDEGSLLHRLALRDEKTSKVENLAASLPPEFCARHGRKINLASY  
LPPGLALRPLEQWAAVGVSPHVRHGLTKNLCEVADLVSLVHADTLPPLAWHRAQK  
DPLSLGDEGLWSPGSQSVWVHFRADQRIIRFLQVCPAGAGALEPFGPSCYL  
DAGLRRLEEGVSCWTLLOAPGRAVLVPAGAPHQVOGLVSTVSTQHLSPETSAL  
SAQLCHQGPSLPPDCHLLAYQMDWAVFQAVKAVGTLOEAK"

## ORIGIN

Query Match 88.0%; Score 2817.4; DB 9; Length 3564;

Best Local Similarity 97.1%; Pred. No. 0;

Matches 2869; Conservative 0; Mismatches 86; Indels 0; Gaps 0;

QY	1	TTTTACTACAGGATCCGAGCATTCACAGTTGGCAAGGAGCCCTTGGCAGCTCGGAA	60
DB	610	TTTTACTACAGGATCCGAGCATTCACAGTTGGCAAGGAGCCCTTGGCAGCTCGGAA	669
QY	61	CCTGGGTTGTTGGCTTAAACTCTGTGGGCACCTGCAGAGAGCCGGGAGGCCGAACGC	120
DB	670	CCTGGGTTGTTGGCTTAAACTCTGTGGGCACCTGCAGAGAGCCGGGAGGCCGAACGC	729
QY	121	CCTTCACTGCACAGAGGATGAGAGATGGGAGCTGGCGGAGCAGAGATCCTTGCCTG	180
DB	730	CCTTCACTGCACAGAGGATGAGAGATGGGAGCTGGCGGAGCAGAGATCCTTGCCTG	789
QY	181	CTCTTCTGGGCGAGCAGACACTGTGCCCTGACCTCTCTGGCCCTCTCTGCCCGCAGGC	240
DB	790	CTCTTCTGGGCGAGCAGACACTGTGCCCTGACCTCTCTGGCCCTCTCTGCCCGCAGGC	849
QY	241	CTTGTTCATACCTCTGGGCAAGCTCTGGGCTGGGCCAGGCGATGGGAACCTTGGGTACCAG	300
DB	850	CTTGTTCATACCTCTGGGCAAGCTCTGGGCTGGGCCAGGCGATGGGAACCTTGGGTACCAG	909
QY	301	CTGGGCGCACAGCAACACAGAGTGCCTCTCTCTAGAGCGCTGTACCCAGCGGGC	360
DB	910	CTGGGCGCACAGCAACACAGAGTGCCTCTCTCTAGAGCGCTGTACCCAGCGGGC	969
QY	361	TGCTGTTCATCTACCCACCACTAAAGTGGGATCTTGGCCCTTGTGGGAAGTGCAG	420
DB	970	TGCTGTTCATCTACCCACCACTAAAGTGGGATCTTGGCCCTTGTGGGAAGTGCAG	1029
QY	421	GAGGCGCTGGAGGGGGTGCCAGTGGAGCCAGCGAACCACAGCGAGGAAGTGAACAAGGCC	480
DB	1030	GAGGCGCTGGAGGGGGTGCCAGTGGAGCCAGCGAACCACAGCGAGGAAGTGAACAAGGCC	1089
QY	481	TTTGGGCCACAGGCGCTGTCCCGCCAGCCACACACAGCTGAAGAAGCATGGCTACA	540
DB	1090	TTTGGGCCACAGGCGCTGTCCCGCCAGCCACACACAGCTGAAGAAGCATGGCTACA	1149
QY	541	CGGCACTCGAGAGAGTTTGAATGTCACGCGGCTGCCCTGAGGTTCAGAGAGAGCGGTT	600
DB	1150	CGGCACTCGAGAGAGTTTGAATGTCACGCGGCTGCCCTGAGGTTCAGAGAGAGCGGTT	1209
QY	601	GCTTCGGCTCCGGGCGCTCAAAAGGCGCAGGAGCCCGGAGTCCAGGGAGCAATGGGCAGT	660
DB	1210	GCTTCGGCTCCGGGCGCTCAAAAGGCGCAGGAGCCCGGAGTCCAGGGAGCAATGGGCAGT	1269
QY	661	CCAGCCCGCAAGGCGCCACCGGACCTTTCCAGGCGATCTCAGAACGAGGGGCTGGGGT	720
DB	1270	CCAGCCCGCAAGGCGCCACCGGACCTTTCCAGGCGATCTCAGAACGAGGGGCTGGGGT	1329
QY	721	TGCGAGAGGTTGCGGACACATCATAGGAAACAAGATGTGACCTCGGACAGCATGAT	780
DB	1330	TGCGAGAGGTTGCGGACACATCATAGGAAACAAGATGTGACCTCGGACAGCATGAT	1389

QY	781	GAGCGAAAGGAGCCCAAGATGGCGAGCCAGTCTTCAGGACCCGGGACTTCAGGACATA	840
DB	1390	GAGCGAGAGGAGCCCGAGACGGCCAGGCGAGTCTTCAGGACCCAGGACTTCAGGAAATA	1449
QY	841	CCATGCTGCTTCTCCCTGCAAACTGGGCTCAATGCTCAAGTGTGCCAGGACGCTGA	900
DB	1450	CCATGCTGCTTCTCCCTGCAAACTGGGCTCAATGCTCAAGTGTGCCAGGACGCTGA	1509
QY	901	GAGGAGGAGGAGCGCTCCACTCTCAGCAAGTGGCGAGATCGCTCTTGGGAGGGAG	960
DB	1510	GAGGAGGAGGAGCGCTCCACTCTCAGCAAGTGGCGAGATCGCTCTTGGGAGGGAG	1569
QY	961	CTGCAGCAGGAGGAGACACAGCCACCAACTCCAGCTCTGAGGAAAGCCAGGCTCGGC	1020
DB	1570	CTGCAGCAGGAGGAGACACAGCCCGCAACTCCAGCTCTGAGGAAAGCCAGGCTCGGC	1629
QY	1021	CTTGACAGCGGCTCAGCAGAGCCCTCGCGCAGCACCTCTCAGTGTGTGGGGACCGA	1080
DB	1630	CTTGATGGCAGGCTCAGCAGAGCCCTCTCCAGCACCTCTCAGTGTGTGGGGACCGA	1689
QY	1081	CTGTGCGGCTCTGTGGGAGGAGCGGAGGCGCTTGGGCTTGGGCCCCAGCGGGAAGCCAA	1140
DB	1690	CTGTGCGGCTCTGTGGGAGAGCGGGAAGCCCTGGCTGGGCCCCAGCGGGAAGCCAA	1749
QY	1141	GGGCCAGCGCTGACAGGGGACGCCAGGATTTCAAGCTGTCTGAGAGCGTTCGCCACAT	1200
DB	1750	GGGCCAGCGCTGACAGGGGACGCCAGGATTTCAAGCTGTCTGAGAGCGTTCGCCACAT	1809
QY	1201	GGACTCTTCAACACCACTGGCGATGTCCCGCTGCAGCACCGGCTGTGTGGGCTGT	1260
DB	1810	GGACTCTTCAACACCACTGGCGATGTCCCGCTGCAGCACCGGCTGTGTGGGCTGT	1869
QY	1261	GGTGTGTGGCAGGCACTGGGGGGCCAGGGAGAAAGCAGGCTTTTCAGGAGCAGTCCGG	1320
DB	1870	GGTGTGTGGCAGGCGCTGGCGGGCCAGGAGAAAGCAGGCTTTCGGGAGCAGTCCAG	1929
QY	1321	GAGGATGACGAGAGCGCGGCGACGCTCTTCCCTGTTCCTGTGATGTGACCCAGTTTGT	1380
DB	1930	GAGGATGACGAGAGCGCGGCGACGCTCTTCCCTGTTCCTGTGATGTGACCCAGTTTGT	1989
QY	1381	TCCAGCCAGGCTTTGCGAGAGCTGAGCACTGCATACACAGCTCTGGGCTCAAGTTTGT	1440
DB	1990	TCCAGCCAGGCTTTGCGAGAGCTGAGCACTGCATACACAGCTCTGGGCTCAAGTTTGT	2049
QY	1441	ATCCGGGGGCACTGCCCTGCCAAGCTGATGCCCGGCTATGGGCCCGCGGGATGCAGGC	1500
DB	2050	ATCCGGGGGCACTGCCCTGCCAAGCTGATGCCCGGCTATGGGCCCGCGGGATGCAGGC	2109
QY	1501	CAGCGAAGGAATCAACAGAAAACGCCCCAACTCCACAACTTCTCTGCAATGGCGAC	1560
DB	2110	CAGCGAAGGAATCAACAGAAAACGCCCCAACTCCACAACTTCTCTGCAATGGCGAC	2169
QY	1561	ACCCAGAGACCAAGAGCATCAAAAGAGGAGACCCCGATTTCGCTGAGACCCAGAGAG	1620
DB	2170	ACCCAGAGACCAAGAGCATCAAAAGAGGAGACCCCGATTTCGCTGAGACCCAGAGAG	2229
QY	1621	GACCGTGTGGCGAGGGCCCTTGCCTTGTCTTCTCTGCGAACTGTGGCTTCTACC	1680
DB	2230	GACCGTGTGGCGAGCGCCCTTGCCTTGTCTTCTCTGCGAACTGTGGCTTCTACC	2289
QY	1681	GGGCTCAAACTCTGTGGGCGCATGAGCGAATACATGCGCCTTGCSCCCCTCACTCCG	1740
DB	2290	GGGCTCAAACTCTGTGGGCGCATGAGCGAATACATGCGCCTTGCSCCCCTCACTCCG	2349
QY	1741	GCCTGCCAGTATGACCGCATCAACATCCTGGACAGCATTTATCGCACAGGTGGTG	1800
DB	2350	GCCTGCCAGTATGACCGCATCAACATCCTGGACAGCATTTATCGCACAGGTGGTG	2409
QY	1801	GAAACGGAAGTCCAGGAGAAAGCCCTGGGCGCGGGCTTCGAGCTGCGCGGGTCTGCGC	1860
DB	2410	GAAACGGAAGTCCAGGAGAAAGCCCTGGGCGCGGGCTTCGAGCTGCGCGGGTCTGCGC	2469
QY	1861	AAGGCTTGGGCTGTGCCCCCTCTCTCAGTGGGGCCCGGCTGCTCCCCCAGGGGCTTG	1920

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 29, 2004, 11:46:34 ; Search time 1179 Seconds  
(without alignments)  
11537.515 Million cell updates/sec

Title: US-10-024-368-1  
Perfect score: 3202  
Sequence: 1 ttttactacaaggtccgag.....tggtcccttttatgtacagg 3202

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : N Geneseq 29Jan04: \*  
1: geneseqn1980s: \*  
2: geneseqn1990s: \*  
3: geneseqn2000s: \*  
4: geneseqn2001as: \*  
5: geneseqn2001bs: \*  
6: geneseqn2002as: \*  
7: geneseqn2003as: \*  
8: geneseqn2003bs: \*  
9: geneseqn2003cs: \*  
10: geneseqn2004s: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3202	100.0	3202	6	AAD31072 Alternati
2	3192.4	99.7	3202	6	AAD31077 Human Hai
3	3190.8	99.7	3202	8	ACD26351 Human Hai
4	3185	99.5	5709	7	ABX13987 Human hai
5	2937.6	91.7	3567	2	Aaz06349 Human Hai
6	533.2	16.7	538	7	ACA57497 Human adi
c 7	86.8	2.7	335	6	ABs69028 Novel mur
8	66	2.1	429	7	ABx41049 Bovine ES
9	62.4	1.9	114955	2	AAx53491 Human ade
c 10	60	1.9	2000	7	ADA71938 Rice gene
11	56.2	1.8	2000	7	ADA71938 Rice gene
12	55.4	1.7	65	6	ABN52357 Mouse spl
13	54.6	1.7	4783	3	AAc76742 Human ORF
c 14	52	1.6	1402	2	AAV36085 DNA encod
15	50	1.6	10732	3	AAa10594 Gene enco
16	49	1.5	125401	4	AAa10594 Gene enco
17	48.2	1.5	427	5	AAf67464 Novel hum
18	48.2	1.5	479	8	ACH34443 Human end
19	48.2	1.5	2359	4	AAi61114 Human pol
20	48.2	1.5	3044	6	ABQ54675 Human ova
21	48.2	1.5	4628	4	AAi59328 Human pol
22	48.2	1.5	6204	5	ABv27441 Human pro
23	48.2	1.5	6204	5	ABv21621 Human pro

24	48.2	1.5	6995	10	ADE40508
c 25	48.2	1.5	110000	4	AAI99682_39
c 26	48	1.5	114955	2	AAx53491
27	47.8	1.5	1089	2	AAZ32065
28	47.8	1.5	1089	5	AAc90322
29	47.6	1.5	6741	3	AAa10595
30	47.4	1.5	339	2	AAQ03321
31	47.4	1.5	747	2	AAQ03323
32	47.4	1.5	954	1	AAa90505
33	47.4	1.5	957	2	AAV66836
34	47.4	1.5	957	2	AAZ32063
35	47.4	1.5	957	5	AAc90320
36	47.4	1.5	1089	1	AAa92576
37	47.4	1.5	1089	1	AAa92576
38	47.2	1.5	2386	2	AAQ63477
39	46.2	1.4	1300	6	ABK83980
40	46.2	1.4	4309	6	ABV77882
41	46.2	1.4	4309	9	ADD18798
42	46.2	1.4	4697	4	AAK51593
c 43	46.2	1.4	4746	4	AAK52577
c 44	46.2	1.4	10678	4	AAa105144
c 45	46.2	1.4	10678	4	ABL98032

ALIGNMENTS

RESULT 1  
AAD31072  
ID AAD31072 standard; DNA; 3202 BP.  
XX  
AC AAD31072;  
XX  
DT 18-JUN-2002 (first entry)  
XX  
DE Alternative version of human Hairless gene (HR).  
XX  
KW Human; Hairless protein; Hr protein; HR gene; dermatological condition;  
KW hair loss; gene therapy; ds.  
XX  
OS Homo sapiens.  
XX  
FH Key  
CDS  
FT  
FT  
FT  
FT  
FT  
FT  
XX  
US6348348-B1.  
XX  
PD 19-FEB-2002.  
XX  
PF 07-APR-1999; 99US-00287354.  
XX  
PR 07-APR-1998; 98US-0080889P.  
XX  
PA (CARN-) CARNEGIE INST WASHINGTON.  
XX  
PI Thompson CC;  
XX  
DR WPI; 2002-204622/26.  
XX  
PT P-PSDB; AAE19794.  
XX  
PT Novel expression construct, useful in the diagnosis and treatment of  
PT dermatological conditions, such as hair loss, contains a Hairless gene  
sequence.  
XX  
PS Claim 4; Col 35-38; 48pp; English.  
XX  
CC The invention relates to human Hairless (Hr) polypeptides and nucleic  
CC acid molecules (HR) encoding such polypeptides. The invention also







QY 769 YLPLGLRLPLEPOLMAAYGVSPHRHGLTKNLVCEVADLVLSILVHADTPLPAWHRAQKD 828  
Db 478 YLPLGLTLHLPLEPOLMAAYGVNSHRHGLTKNLVCEVSDLSILVHAEAPLPWYRAQKD 537  
QY 829 FLSGDGEGLWSPGSOYSTVWHVFRADQARIRFLQWVCPAGAGALEPGAGSCYLDAG 888  
Db 538 FLSGDGEGLWSPGSOYSTVWHVFRADQARIRFLQWVCPAGAGALEPGAGSCYLDAG 597  
QY 889 LRRRLREEWGYSCTLLQAPGEALVLPAGAPHQVGLVSTVTOHFLSPETSALSACL 948  
Db 598 LRRRLREEWGYSCTLLQAPGEALVLPAGAPHQVGLVSTVTOHFLSPETSALSACL 657  
QY 949 HQGSLPPDCHLLVYAQMDWAVFQAVKAVGTIOLEAK 984  
Db 658 HQGSLPPDCHLLVYAQMDWAVFQAVKAVGTIOLEAK 693

RESULT 8  
US-10-408-765A-795  
; Sequence 795, Application US/10408765A  
; Publication No. US20040101874A1  
; GENERAL INFORMATION:  
; APPLICANT: Ghosh, Soumitra S.  
; APPLICANT: Faby, Eoin D.  
; APPLICANT: Zhang, Bing  
; APPLICANT: Gibson, Bradford W.  
; APPLICANT: Taylor, Steven W.  
; APPLICANT: Glenn, Gary M.  
; APPLICANT: Warnock, Dale E.  
; TITLE OF INVENTION: TARGETS FOR THERAPEUTIC INTERVENTION  
; FILE REFERENCE: 660088.465  
; CURRENT APPLICATION NUMBER: US/10/408,765A  
; CURRENT FILING DATE: 2003-04-04  
; NUMBER OF SEQ ID NOS: 3077  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 795  
; LENGTH: 1265  
; TYPE: PRF  
; ORGANISM: Homo sapiens  
US-10-408-765A-795

Query Match 14.5%; Score 778; DB 16; Length 1265;  
Best Local Similarity 26.9%; Pred. No. 3.5e-46;  
Matches 280; Conservative 144; Mismatches 360; Indels 259; Gaps 40;

QY 157 VNKAS-----GPRACPPSHHTKKTWLRHSEQ-----FECPRGC--- 192  
Db 253 INKANSVNGQASQTSQPNYHTKLKAWLTHSEEDKNTNMKNSGNSVSEIHKFCSVNL 312  
QY 193 -----PEVEERPVRLALK-----RAGSEPVQAMGSPAPKRP 227  
Db 313 LASTSSDIQNSVDSKIIVDKYVKKDKVNRKAKRTYESGSDGSDSESKSEQRTKRP 372  
QY 228 DPF-----PCTAEOGAGCHQEVDTSIGN-----KVDSSQHQDEQ 262  
Db 373 KPTYKKQNDLQKRKETEEDLKNPGLVSRSAKRSKLQNSNTGTPRSVLKDWKVKY 432  
QY 263 KGPQDQASLDQPGLDIPLALPAKLAQCSQ-----AQAGGGGHA-----CHSQQVRR-- 314  
Db 433 KUKQTGESFLQDS-----CCSIGPNLQKRCRLIRSKKGEPAHSPVFCFFYFRRLS 487  
QY 315 -----SPLGELQBEETATNSSSEEGSGDPSRLSTGTALHLLSLGLDRL 361  
Db 488 FSKNGVVRIDGFSPP-----DQVDEAMSLWTHE---NFEDELDIETSKYILDIIGK 539  
QY 362 CILLRERERALAWAQREGO---GPATVGDSPGIPRCCSCHGLFNTHWRCPCHRLCV 418  
Db 540 COLVTSKALSWVKDKAKIAWKAVR-----GVREMCDAEATLFNIHWVCKGCFVCL 595  
QY 419 AGRVAGTGRAREKAGFQBSAECTQEAAGAACSLMLTQFVSSQALAEELSTAMHGVVK 478

Db 596 DCYKAKERKSRDX---ELYANMKCVKQPHDHHKMLPTQIIPGSLVTLDDLDMHTLREK 652  
QY 479 FDIRGHCP-----QADARVWAPGDA 499  
Db 653 YGKSHCHCTNKQNLQVGNFPTMNGSVQLQNLVHNSNKISLCPESQQQNTPPKSEKNG 712  
QY 500 QOQKES-----TQKTPTPPQSC-----NGDTHRTKSIKEE--TPDSAE 536  
Db 713 GSPESVDGTNKTUTPESQSPHLWLADLAQKAREKKEKKEJLTLENQIKEEREQNSE 772  
QY 537 TPAEDRAGR--GPLPC-----PSICELLASTAVKLCIGHERIHMAFAPV-----TPALPSD 585  
Db 773 SP-----NGRTSPVSNQNEQSTLRDLTTTAGKLRVGSTDAIAFAFVYVMGAPSSKSG 828  
QY 586 DRINILDSIIAQQVVERKIOEKA-----LGPLGRAP-----GLRGLGLPLSPVR 631  
Db 829 RTMPNILDIIASVVENKIPPSKTSKINVKPELKEEPEESIISAVENNKLYSDIPHSWIC 888  
QY 632 PRLPPGALLWLOBPQPCPRGFHLFOHWHQOGOPVLVSGIQTLOGNLWGTALGALGG 691  
Db 889 EK-----HILWLKDYK--NSSNWKLKFKCKWKQOGQPAVVGVHKKNNISLWKAESISLDFG 941  
QY 692 QVQA--LSPLGPPQPSLSGTTTFWEGFSWPELRPKSDEGSVLLHLRALGD---ED----- 741  
Db 942 DHQADLLNCKDSIISNANVKEFDGFEVSKRKNKSGETVVL--KLKDWPSGEDFKTMM 999  
QY 742 TSVENLAASLPPEYCALHOKLNIASVLPGLALRP--LEPOLWAAVGVSPHRGH--LGTK 799  
Db 1000 PARYEDLLKSLPEYCNPEGKFNASHL--PGFVRPDLGPRLCGAYGVAAKDHIDGTT 1058  
QY 800 NLCVEVADLVSIHADTPLPAWHRAQKDFLSGLDGEGL-----WSPGSQV--STVWHV 851  
Db 1059 NHHIEVSDVNILVYVGIAGKNGILSKAGILKKEEEDLDLILKRLKDSSEIPGALWHI 1118  
QY 852 FRAQDAQRIIRFLQWVC--PAGAGAL---EPGAPGSCYLDAGLRRLREEWGYSCTWLLQA 907  
Db 1119 YAGRDVDKIREFLQKISKEQGLEVLPEHDPTRDQSWYVKNKLRQLLEEGYVRTCTLIQF 1178  
QY 908 PGEAVLPAGAPHQVQGLVSTVTOHFLSPETSALSACLCHQGPSLPPDCHLLVYAQMDW 967  
Db 1179 LGDAIVLPAGALHQVQNFHSCIQVTEDFVSPHILVESPHLTQ-----ELRLKKEINY 1231  
QY 968 -----AVFQAVKAVGTIO 981  
Db 1232 DDKLQVKNILYHAKEMVRALK 1253

RESULT 9  
US-10-193-874-20  
; Sequence 20, Application US/10193874  
; Publication No. US20030064396A1  
; GENERAL INFORMATION:  
; APPLICANT: Jenapharm GmbH & Co. KG  
; TITLE OF INVENTION: Comodulators of Nuclear Receptors  
; FILE REFERENCE: 52145  
; CURRENT APPLICATION NUMBER: US/10/193,874  
; CURRENT FILING DATE: 1998-11-18  
; NUMBER OF SEQ ID NOS: 20  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 20  
; LENGTH: 2055  
; TYPE: PRF  
; ORGANISM: Homo sapiens  
US-10-193-874-20

Query Match 14.5%; Score 777.5; DB 12; Length 2055;  
Best Local Similarity 26.8%; Pred. No. 7.3e-46;  
Matches 280; Conservative 144; Mismatches 360; Indels 259; Gaps 40;

QY 157 VNKAS-----GPRACPPSHHTKKTWLRHSEQ-----FECPRGC--- 192  
Db 1042 INKANSVNGQASQTSQPNYHTKLKAWLTHSEEDKNTNMKNSGNSVSEIHKFCSVNL 1101



Db 506 EGGHACHSQVRRSPGLGELQOEEDTATNSSSEEGPGSGPDSRLSTGLAKHLLSGLGDR 565  
Qy 361 LCRLLRERREALAWAQREGQGPVATGSPGIPRCCSRCHHGLFNTHWRCPSHRLCVAC 420  
Db 566 LCRLLRERREALAWAQREGQGPVATGSPGIPRCCSRCHHGLFNTHWRCPSHRLCVAC 625  
Qy 421 GRVAGTGRAREKAGFOQSAEECTQEAAGAACSLMLTQFVSSQALAEIATAMHVVWVKFD 480  
Db 626 GRVAGTGRAREKAGFOQSAEECTQEAAGAACSLMLTQFVSSQALAEIATAMHVVWVKFD 685  
Qy 481 IRGHCPCQADARWAPADGAGQKESTOKTPTPOPCNGDTHRTKSIKEETPDGSAETPAE 540  
Db 586 IRGHCPCQADARWAPADGAGQKESTOKTPTPOPCNGDTHRTKSIKEETPDGSAETPAE 745  
Qy 541 DRAGKGLPCPSLCELLASTAVKLCGLHERIHMFAFVTPALPDDRTITNLDSTIIAQVV 600  
Db 746 DRAGKGLPCPSLCELLASTAVKLCGLHERIHMFAFVTPALPDDRTITNLDSTIIAQVV 805  
Qy 601 ERKIQEALGPGLRAGPGLRGLGLPLSPVRPRPLPPGALLWLOEPQCPRRGHFLFOEH 660  
Db 806 ERKIQEALGPGLRAGPGLRGLGLPLSPVRPRPLPPGALLWLOEPQCPRRGHFLFOEH 865  
Qy 661 WRQGPVLVSGIQTLOGLNMGTEALGALGOVQALSPGPPQPSLSGSTTFWEGFSWPE 720  
Db 866 WRQGPVLVSGIQTLOGLNMGTEALGALGOVQALSPGPPQPSLSGSTTFWEGFSWPE 925  
Qy 721 LRPKSDEGSVLLHRAFGDEDTSRVENLAASLPLPEYCALHGKLNLSYLPGLALRPLE 780  
Db 926 LRPKSDEGSVLLHRAFGDEDTSRVENLAASLPLPEYCALHGKLNLSYLPGLALRPLE 985  
Qy 781 POLMAAYGVSPHRGHGLTKNLCEVADIVSLVHADTPLPAWHRAQKDFLSGLDGEGLWS 840  
Db 986 POLMAAYGVSPHRGHGLTKNLCEVADIVSLVHADTPLPAWHRAQKDFLSGLDGEGLWS 1045  
Qy 841 PGSQVSTVWHVFRADQAOIRRFLOMVCPCAGALEPGAGSCYLDAGLRRLREEWGVS 900  
Db 1046 PGSQVSTVWHVFRADQAOIRRFLOMVCPCAGALEPGAGSCYLDAGLRRLREEWGVS 1105  
Qy 901 CWTLLQAPGEAVLPAGAPHVQGLVSTVSTQHFPLSPETSALSQALCHOGFSLPPDCHL 960  
Db 1106 CWTLLQAPGEAVLPAGAPHVQGLVSTVSTQHFPLSPETSALSQALCHOGFSLPPDCHL 1165  
Qy 961 LYAQMWAFAVFAVKVAVGTLOEAK 984  
Db 1166 LYAQMWAFAVFAVKVAVGTLOEAK 1189

RESULT 5  
US-10-024-368-6  
; Sequence 6, Application US/10024368  
; Publication No. US20030027300A1  
; GENERAL INFORMATION:  
; APPLICANT: THOMPSON, Catherine C.  
; TITLE OF INVENTION: HUMAN HAIRLESS GENE AND PROTEIN  
; FILE REFERENCE: Thompson-20263/0243435  
; CURRENT APPLICATION NUMBER: US/10/024,368  
; CURRENT FILING DATE: 2001-12-21  
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US/09/287,354  
; PRIOR FILING DATE: EARLIER FILING DATE: 1999-04-07  
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US 60/080,888  
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-04-07  
; NUMBER OF SEQ ID NOS: 10  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 6  
; LENGTH: 1182  
; TYPE: PRT  
; ORGANISM: Mouse  
US-10-024-368-6

Query Match 76.6%; Score 4102.5; DB 14; Length 1182;  
Best Local Similarity 78.4%; Pred. No. 1.8e-282;  
Matches 771; Conservative 62; Mismatches 144; Indels 7; Gaps 5;

Qy 1 FYKDPSPRLAKEPLAAAPGLGLNSGGHLOQAGEAERPSLHORDEMGAGQOQNPCP 60  
Db 206 FYKDPSPRLAKEPL--ABSGMLGLAPGGHLOQACESEGLSHQRDGETAGGQQNLCP 263  
Qy 61 LFLGQPTVPTWTSWPAACPPGLVHTLGNWAGPGDGNLGYQLGPPATPRCPSPEPVTQRG 120  
Db 264 VFLGYPTVTPRAPWPSCPPGLVHSLGNIWAGPSNSLGYQLGPPATPRCPSPPPTPGG 323  
Qy 121 CCSSYPTKGGDLGPKCKCKEGLEGASGASBPSEEVNKGASGPPACPPSHHTKLUKKTWLT 180  
Db 324 CCSSHLPAEGDLGPKCKCKDSDPEGGSGPESSEERNKADS-RACPPSHHTKLUKKTWLT 382  
Qy 181 RKSQFPCPCPCPEVERPVARLPAKAKRAGSPVQAGMSPAPKPPDPPTGTAEOQAGG 240  
Db 383 RHSEQFCPCGCSCKEBSATGLRALKRAGSPVQAGMSPAPKPPDPPTGTAEOQAGG 442  
Qy 241 WQEVDRTSIGNKVDSDQDEKQPOGQASLQDIPCLALPAKLAQCCQCAQAAG 300  
Db 443 WQETPETIIGSK-AEAQOEORQPRDRIRLOESRLVDTSCQHLLAGVTCCQSCVQAAG 501  
Qy 301 EGGHACHSQVRRSPGLGELQOEEDTATNSSSEEGPGSGPDSRLSTGLAKHLLSGLGDR 360  
Db 502 EVGLTGHKSQKRRSPL-EEKQLEEDSSATSEEGGGPGPEASLNGKLAKHLLSGLGDR 560  
Qy 361 LCRLLRERREALAWAQREGQGPVATGSPGIPRCCSRCHHGLFNTHWRCPSHRLCVAC 420  
Db 561 LCRLLRERREALAWAQREGQGPVATGSPGIPRCCSRCHHGLFNTHWRCPSHRLCVAC 620  
Qy 421 GRVAGTGRAREKAGFOQSAEECTQEAAGAACSLMLTQFVSSQALAEIATAMHVVWVKFD 480  
Db 621 GRIAGAKNREKTSQEQHTDDCAQEAAGAACSLMLTQFVSSQALAEIATAMHVVWVKFD 680  
Qy 481 IRGHCPCQADARWAPADGAGQKESTOKTPTPOPCNGDTHRTKSIKEETPDGSAETPAE 540  
Db 681 IRGHCPCQADARWAPADGAGQKESTOKTPTPOPCNGDTHRTKSIKEETPDGSAETPAE 740  
Qy 541 DRAGKGLPCPSLCELLASTAVKLCGLHERIHMFAFVTPALPDDRTITNLDSTIIAQVV 600  
Db 741 DRAGKGLPCPSLCELLASTAVKLCGLHERIHMFAFVTPALPDDRTITNLDSTIIAQVV 800  
Qy 601 ERKIQEALGPGLRAGPGLRGLGLPLSPVRPRPLPPGALLWLOEPQCPRRGHFLFOEH 660  
Db 801 ERKIQEALGPGLRAGPGLRGLGLPLSPVRPRPLPPGALLWLOEPQCPRRGHFLFOEH 858  
Qy 661 WRQGPVLVSGIQTLOGLNMGTEALGALGOVQALSPGPPQPSLSGSTTFWEGFSWPE 720  
Db 859 WRQGPVLVSGIQTLOGLNMGTEALGALGOVQALSPGPPQPSLSGSTTFWEGFSWPE 918  
Qy 721 LRPKSDEGSVLLHRAFGDEDTSRVENLAASLPLPEYCALHGKLNLSYLPGLALRPLE 780  
Db 919 TRPKLDEGSVLLHRTLGDKDASRVQNLAASLPLPEYCALHGKLNLSYLPGLALRPLE 978  
Qy 781 POLMAAYGVSPHRGHGLTKNLCEVADIVSLVHADTPLPAWHRAQKDFLSGLDGEGLWS 840  
Db 979 POLMAAYGVSPHRGHGLTKNLCEVADIVSLVHADTPLPAWHRAQKDFLSGLDGEGLWS 1038  
Qy 841 PGSQVSTVWHVFRADQAOIRRFLOMVCPCAGALEPGAGSCYLDAGLRRLREEWGVS 900  
Db 1039 PGSQVSTVWHVFRADQAOIRRFLOMVCPCAGALEPGAGSCYLDAGLRRLREEWGVS 1098  
Qy 901 CWTLLQAPGEAVLPAGAPHVQGLVSTVSTQHFPLSPETSALSQALCHOGFSLPPDCHL 960  
Db 1099 CWTLLQAPGEAVLPAGAPHVQGLVSTVSTQHFPLSPETSALSQALCHOGFSLPPDCHL 1158  
Qy 961 LYAQMWAFAVFAVKVAVGTLOEAK 984  
Db 1159 LYAQMWAFAVFAVKVAVGTLOEAK 1182

RESULT 6  
US-10-024-368-5  
; Sequence 5, Application US/10024368  
; Publication No. US20030027300A1

QY 961 LYAQMDFVFAVQVAVGTLQEA 984  
Db 1166 LYAQMDFVFAVQVAVGTLQEA 1189

RESULT 3  
US-10-122-013-17  
; Sequence 17, Application US/10122013  
; Publication No. US20030077614A1  
; GENERAL INFORMATION:  
; APPLICANT: Cristiano, Angela  
; TITLE OF INVENTION: NUCLEIC ACIDS FOR INHIBITING HAIRLESS PROTEIN EXPRESSION AND METH  
; FILE OF INVENTION: USE THEREOF  
; FILE REFERENCE: 0575/62637A  
; CURRENT APPLICATION NUMBER: US/10/122,013  
; CURRENT FILING DATE: 2002-04-12  
; NUMBER OF SEQ ID NOS: 27  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 17  
; LENGTH: 1189  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-122-013-17

Query Match 99.5%; Score 5331; DB 14; Length 1189;  
Best Local Similarity 99.7%; Pred. No. 0;  
Matches 981; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 FFKPSPILRLAKEPLAAAEPLGLNLSGGHLOAGEAERPSLHQRDGMGAGRQNP 60  
Db 206 FFKPSPILRLAKEPLAAAEPLGLNLSGGHLOAGEAERPSLHQRDGMGAGRQNP 265

QY 61 LFLGPDTPVWTSWPAACPPGLVHTLGNVWAGPDGNLGYQLGPPATPCPSPEPPVTQ 120  
Db 266 LFLGPDTPVWTSWPAACPPGLVHTLGNVWAGPDGNLGYQLGPPATPCPSPEPPVTQ 325

QY 121 CCSSYPPTKGGDLGPCCKQEGLEGASGASFPSEEVNKAASPRCPSSHHTKLKXTWLT 180  
Db 326 CCSSYPPTKGGDLGPCCKQEGLEGASGASFPSEEVNKAASPRCPSSHHTKLKXTWLT 385

QY 181 RHSEGFECPRGCPVEERVARLRALKRAGSEVQVAMGSPAPKRPDPFPGTAEGAGG 240  
Db 386 RHSEGFECPRGCPVEERVARLRALKRAGSEVQVAMGSPAPKRPDPFPGTAEGAGG 445

QY 241 WQEVDTISGNKDVSGQDEKQPGDQASLQDPGLQDIPCLALPAKLAQCQSCAAAG 300  
Db 446 LQEVDTISGNKDVSGQDEKQPGDQASLQDPGLQDIPCLALPAKLAQCQSCAAAG 505

QY 301 EGGHACHSQVRRSPGLGELQOEEDTATNSSSEEGPGSPDSRLSTGLAKHLLSLGLDR 360  
Db 506 EGGHACHSQVRRSPGLGELQOEEDTATNSSSEEGPGSPDSRLSTGLAKHLLSLGLDR 565

QY 361 LCLLRERREALAWAQREGQAVTGDSPGIPRCCSRCHHGLFNTHWRCPRSHRLCVAC 420  
Db 566 LCLLRERREALAWAQREGQAVTGDSPGIPRCCSRCHHGLFNTHWRCPRSHRLCVAC 625

QY 421 GRVAGTGRAREKAGFOEQSAEECTQAGHAACSLMLTQFVSSQALAEELSTAMHVVWVFD 480  
Db 626 GRVAGTGRAREKAGFOEQSAEECTQAGHAACSLMLTQFVSSQALAEELSTAMHVVWVFD 685

QY 481 IGHCFQADARVWAPGDAGQOKESTQKTPPTPQPSGNGDTHRTKSIKEETPDSAPTAE 540  
Db 686 IGHCFQADARVWAPGDAGQOKESTQKTPPTPQPSGNGDTHRTKSIKEETPDSAPTAE 745

QY 541 DRAGGFLPCPSLCELLASTAVKLCIGHERIHWAFAPVTPALPDSDDRTNIILDSIIAQV 600  
Db 746 DRAGGFLPCPSLCELLASTAVKLCIGHERIHWAFAPVTPALPDSDDRTNIILDSIIAQV 805

QY 601 ERKIEKALGPLRAGPLRKLGLPLSVPRPLPPGALLMLQEPQPCPRGGFHLFQEH 660  
Db 806 ERKIEKALGPLRAGPLRKLGLPLSVPRPLPPGALLMLQEPQPCPRGGFHLFQEH 865

QY 661 WRQQPVLVSGIQTTLQGNLWGTEALGALGQVQALSPLPQPPQSSLGSTTFWEGFSWPE 720  
Db 866 WRQQPVLVSGIQTTLQGNLWGTEALGALGQVQALSPLPQPPQSSLGSTTFWEGFSWPE 925

QY 721 LRPKSDEGSVLLHRLAGDEDTSRVENLAASLPLPEYCALHGKLNLAASYLPPLGALRPLE 780  
Db 926 LRPKSDEGSVLLHRLAGDEDTSRVENLAASLPLPEYCALHGKLNLAASYLPPLGALRPLE 985

QY 781 POLWAAVGVSPHRGHGTGNLCVEADLVSLVHADTLPFAWHRAOKDFLSGLDGEGLWS 840  
Db 986 POLWAAVGVSPHRGHGTGNLCVEADLVSLVHADTLPFAWHRAOKDFLSGLDGEGLWS 1045

QY 841 PGSQVSTVWHVFAQAQRIIRFLQMVCPAGAGALEPFGPGSCYLDAGLRRLRREWGV 900  
Db 1046 PGSQVSTVWHVFAQAQRIIRFLQMVCPAGAGALEPFGPGSCYLDAGLRRLRREWGV 1105

QY 901 CWTLLQAPGEAVLVPAGAPHQVGLVSTVSTVTOHFLSPETSALSAQICHQGPSLPDCHL 960  
Db 1106 CWTLLQAPGEAVLVPAGAPHQVGLVSTVSTVTOHFLSPETSALSAQICHQGPSLPDCHL 1165

QY 961 LYAQMDFVFAVQVAVGTLQEA 984  
Db 1166 LYAQMDFVFAVQVAVGTLQEA 1189

RESULT 4  
US-10-024-368-3  
; Sequence 3, Application US/10024368  
; Publication No. US20030027300A1  
; GENERAL INFORMATION:  
; APPLICANT: THOMPSON, Catherine C.  
; TITLE OF INVENTION: HUMAN HAIRLESS GENE AND PROTEIN  
; FILE REFERENCE: Thompson-20263/0243435  
; CURRENT APPLICATION NUMBER: US/10/024,368  
; CURRENT FILING DATE: 2001-12-21  
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US/09/287,354  
; PRIOR FILING DATE: EARLIER FILING DATE: 1999-04-07  
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US 60/080,888  
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-04-07  
; NUMBER OF SEQ ID NOS: 10  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 3  
; LENGTH: 1189  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-024-368-3

Query Match 98.7%; Score 5290; DB 14; Length 1189;  
Best Local Similarity 99.0%; Pred. No. 0;  
Matches 974; Conservative 1; Mismatches 9; Indels 0; Gaps 0;

QY 1 FFKPSPILRLAKEPLAAAEPLGLNLSGGHLOAGEAERPSLHQRDGMGAGRQNP 60  
Db 206 FFKPSPILRLAKEPLAAAEPLGLNLSGGHLOAGEAERPSLHQRDGMGAGRQNP 265

QY 61 LFLGPDTPVWTSWPAACPPGLVHTLGNVWAGPDGNLGYQLGPPATPCPSPEPPVTQ 120  
Db 266 LFLGPDTPVWTSWPAACPPGLVHTLGNVWAGPDGNLGYQLGPPATPCPSPEPPVTQ 325

QY 121 CCSSYPPTKGGDLGPCCKQEGLEGASGASFPSEEVNKAASPRCPSSHHTKLKXTWLT 180  
Db 326 CCSSYPPTKGGDLGPCCKQEGLEGASGASFPSEEVNKAASPRCPSSHHTKLKXTWLT 385

QY 181 RHSEGFECPRGCPVEERVARLRALKRAGSEVQVAMGSPAPKRPDPFPGTAEGAGG 240  
Db 386 RHSEGFECPRGCPVEERVARLRALKRAGSEVQVAMGSPAPKRPDPFPGTAEGAGG 445

QY 241 WQEVDTISGNKDVSGQDEKQPGDQASLQDPGLQDIPCLALPAKLAQCQSCAAAG 300  
Db 446 LQEVDTISGNKDVSGQDEKQPGDQASLQDPGLQDIPCLALPAKLAQCQSCAAAG 505

QY 301 EGGHACHSQVRRSPGLGELQOEEDTATNSSSEEGPGSPDSRLSTGLAKHLLSLGLDR 360  
Db 506 EGGHACHSQVRRSPGLGELQOEEDTATNSSSEEGPGSPDSRLSTGLAKHLLSLGLDR 565

QY 361 LCLLRERREALAWAQREGQAVTGDSPGIPRCCSRCHHGLFNTHWRCPRSHRLCVAC 420  
Db 566 LCLLRERREALAWAQREGQAVTGDSPGIPRCCSRCHHGLFNTHWRCPRSHRLCVAC 625

QY 421 GRVAGTGRAREKAGFOEQSAEECTQAGHAACSLMLTQFVSSQALAEELSTAMHVVWVFD 480  
Db 626 GRVAGTGRAREKAGFOEQSAEECTQAGHAACSLMLTQFVSSQALAEELSTAMHVVWVFD 685

QY 481 IGHCFQADARVWAPGDAGQOKESTQKTPPTPQPSGNGDTHRTKSIKEETPDSAPTAE 540  
Db 686 IGHCFQADARVWAPGDAGQOKESTQKTPPTPQPSGNGDTHRTKSIKEETPDSAPTAE 745

QY 541 DRAGGFLPCPSLCELLASTAVKLCIGHERIHWAFAPVTPALPDSDDRTNIILDSIIAQV 600  
Db 746 DRAGGFLPCPSLCELLASTAVKLCIGHERIHWAFAPVTPALPDSDDRTNIILDSIIAQV 805

QY 601 ERKIEKALGPLRAGPLRKLGLPLSVPRPLPPGALLMLQEPQPCPRGGFHLFQEH 660  
Db 806 ERKIEKALGPLRAGPLRKLGLPLSVPRPLPPGALLMLQEPQPCPRGGFHLFQEH 865

181	QY	RHSEQECPCGCEVERPVARLRAKAGSPVEQAMGSPAPKRPDPDFPGTAEOGAGG	240
181	Db	RHSEQECPCGCEVERPVARLRAKAGSPVEQAMGSPAPKRPDPDFPGTAEOGAGG	240
241	QY	WQVVRTSIGNKVDVDSQHDEKQPGDQASLQDPGLQDIPCLALPAKLAQCSCAQAAQ	300
241	Db	WQVVRTSIGNKVDVDSQHDEKQPGDQASLQDPGLQDIPCLALPAKLAQCSCAQAAQ	300
301	QY	EGGHACHSQOVRRSPLGGELQOEEDTATNSSBEGPGSPDSRLSTGLAKHLSGLGDR	360
301	Db	EGGHACHSQOVRRSPLGGELQOEEDTATNSSBEGPGSPDSRLSTGLAKHLSGLGDR	360
361	QY	LCRLLRERREALAWAQREGQPATVGDSPGTPRCCSRCHHGLFNTHWRCPRCSHRLCVAC	420
361	Db	LCRLLRERREALAWAQREGQPATVGDSPGTPRCCSRCHHGLFNTHWRCPRCSHRLCVAC	420
421	QY	GRVAGTGRAREKAGFQSQABEECTQEAGHAACSLMLTOFVSSQALAEIATMHQVWVKFD	480
421	Db	GRVAGTGRAREKAGFQSQABEECTQEAGHAACSLMLTOFVSSQALAEIATMHQVWVKFD	480
481	QY	IRGHCPQADARVWAPDAGCOQKSGSTQKTPTTPQSPSCNGDTHRTKSIKEETPDSAEETPAE	540
481	Db	IRGHCPQADARVWAPDAGCOQKSGSTQKTPTTPQSPSCNGDTHRTKSIKEETPDSAEETPAE	540
541	QY	DRAGRGLPCPSLCELLASTAVKLCGLGHERIHMAFAPVTPALPSDDRTINILDSIIAQVV	600
541	Db	DRAGRGLPCPSLCELLASTAVKLCGLGHERIHMAFAPVTPALPSDDRTINILDSIIAQVV	600
601	QY	ERKIQEKALGFLRAGPLRKLGLIPLSPVRPRLPPPGALLMLQEPQPCPRRGFHLFOEH	660
601	Db	ERKIQEKALGFLRAGPLRKLGLIPLSPVRPRLPPPGALLMLQEPQPCPRRGFHLFOEH	660
661	QY	WRQCPVLVSGIQRTLOGNLWGTALGALGGQVQALSPLGPPQPSLSGSTTTTWEFGSWPE	720
661	Db	WRQCPVLVSGIQRTLOGNLWGTALGALGGQVQALSPLGPPQPSLSGSTTTTWEFGSWPE	720
721	QY	LRPKSDSGSVLLLRALGDEDTSRVENIAAASLPLPEYCALHGKUNLASVLPPLGALRPLE	780
721	Db	LRPKSDSGSVLLLRALGDEDTSRVENIAAASLPLPEYCALHGKUNLASVLPPLGALRPLE	780
781	QY	PQLWAAAYGVSPHRGHGLTKNLCVEVADLVSIIVHADTTPLPAWHRAQKDFLSGLDGEGLWS	840
781	Db	PQLWAAAYGVSPHRGHGLTKNLCVEVADLVSIIVHADTTPLPAWHRAQKDFLSGLDGEGLWS	840
841	QY	PGSQVSTVWHVFRAGDAQRIRRFLOWCPAGAGALEPGAPGSCYLDAGLRRLRREWGVWS	900
841	Db	PGSQVSTVWHVFRAGDAQRIRRFLOWCPAGAGALEPGAPGSCYLDAGLRRLRREWGVWS	900
901	QY	CWTLLQAPGEAVLPVAPAGPHOVQGLVSTVSVTQHFLSPETSALSACLCHQGSLLPDDCHL	960
901	Db	CWTLLQAPGEAVLPVAPAGPHOVQGLVSTVSVTQHFLSPETSALSACLCHQGSLLPDDCHL	960
961	QY	LYAQMDWAVFOAVKVAGTGTQEAK	984
961	Db	LYAQMDWAVFOAVKVAGTGTQEAK	984

## RESULT 2

US-10-024-368-4  
; Sequence 4, Application US/10024368  
; Publication NO. US20030027300A1  
; GENERAL INFORMATION:  
; APPLICANT: THOMPSON, Catherine C.  
; TITLE OF INVENTION: HUMAN HAIRLESS GENE AND PROTEIN  
; FILE REFERENCE: Thompson-20263/0243435  
; CURRENT APPLICATION NUMBER: US/10/024, 368  
; CURRENT FILING DATE: 2001-12-21  
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US/09/287,354  
; PRIOR FILING DATE: EARLIER FILING DATE: 1999-04-07  
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US 60/080,888  
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-04-07  
; NUMBER OF SEQ ID NOS: 10  
; SOFTWARE: PatentIn Ver. 2.0

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 29, 2004, 10:28:55 ; Search time 54 Seconds  
(without alignments)  
5151.519 Million cell updates/sec

Title: US-10-024-368-2

Perfect score: 5358

Sequence: 1 FYYKDPISIPRLAKEPLAAAE.....MDWAVFQAVKVAVTIQEAK 984

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1166195 seqs, 282705291 residues

Total number of hits satisfying chosen parameters: 1166195

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications, AA:\*

- 1: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep.\*
- 2: /cgn2\_6/ptodata/1/pubpaa/PCT\_NEW\_PUB.pep.\*
- 3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.pep.\*
- 4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pep.\*
- 5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep.\*
- 6: /cgn2\_6/ptodata/1/pubpaa/PCTUS\_PUBCOMB.pep.\*
- 7: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pep.\*
- 8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep.\*
- 9: /cgn2\_6/ptodata/1/pubpaa/US09A\_PUBCOMB.pep.\*
- 10: /cgn2\_6/ptodata/1/pubpaa/US09B\_PUBCOMB.pep.\*
- 11: /cgn2\_6/ptodata/1/pubpaa/US09C\_PUBCOMB.pep.\*
- 12: /cgn2\_6/ptodata/1/pubpaa/US09\_NEW\_PUB.pep.\*
- 13: /cgn2\_6/ptodata/1/pubpaa/US10A\_PUBCOMB.pep.\*
- 14: /cgn2\_6/ptodata/1/pubpaa/US10B\_PUBCOMB.pep.\*
- 15: /cgn2\_6/ptodata/1/pubpaa/US10C\_PUBCOMB.pep.\*
- 16: /cgn2\_6/ptodata/1/pubpaa/US10\_NEW\_PUB.pep.\*
- 17: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pep.\*
- 18: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	5358	100.0	984	14	US-10-024-368-2
2	5343	99.7	1189	14	Sequence 2, Appli
3	5331	99.5	1189	14	Sequence 4, Appli
4	5290	98.7	1189	14	Sequence 17, Appli
5	4102.5	76.6	1182	14	Sequence 3, Appli
6	4067	75.9	1207	14	Sequence 6, Appli
7	3084.5	57.6	693	16	Sequence 5, Appli
8	778	14.5	1265	16	Sequence 93, Appli
9	777.5	14.5	2055	12	Sequence 795, App
10	777.5	14.5	2321	12	Sequence 20, Appli
11	777.5	14.5	2358	12	Sequence 17, Appli
12	777.5	14.5	2540	12	Sequence 16, Appli
13	777.5	14.5	2552	12	Sequence 14, Appli
14	711.5	13.3	1417	12	Sequence 15, Appli
15	696.5	13.0	1212	12	Sequence 19, Appli

16	574	10.7	671	15	US-10-264-049-2730	Sequence 2730, Ap
17	564.5	10.5	953	12	US-10-193-874-18	Sequence 18, Appl
18	287.5	5.4	1057	15	US-10-108-260A-3230	Sequence 3230, Ap
19	286.5	5.3	941	12	US-10-424-599-210074	Sequence 210074,
20	284	5.3	747	12	US-10-425-114-38701	Sequence 38701, A
21	278.5	5.2	837	15	US-10-094-749-2698	Sequence 2698, Ap
22	254.5	4.7	579	12	US-10-425-114-56872	Sequence 56872, A
23	249	4.6	363	12	US-10-425-114-42461	Sequence 42461, A
24	247.5	4.6	541	12	US-10-425-114-72403	Sequence 72403, A
25	247	4.6	357	12	US-10-425-114-69939	Sequence 69939, A
26	223	4.2	349	12	US-10-425-114-74462	Sequence 74462, A
27	223	4.2	380	12	US-10-425-114-71856	Sequence 71856, A
28	205	3.8	1561	16	US-10-408-765A-235	Sequence 235, App
29	203.5	3.8	19652	15	US-10-084-846A-7	Sequence 7, Appli
30	203.5	3.8	19695	15	US-10-084-846A-3	Sequence 3, Appli
31	203	3.8	1386	14	US-10-327-414-2	Sequence 2, Appli
32	200.5	3.7	19608	15	US-10-084-846A-8	Sequence 8, Appli
33	196	3.7	1564	15	US-10-292-798-1254	Sequence 1254, Ap
34	194	3.6	1259	14	US-10-260-715-8	Sequence 8, Appli
35	189	3.5	1497	12	US-10-016-248-8	Sequence 8, Appli
36	186.5	3.5	19725	15	US-10-084-846A-4	Sequence 4, Appli
37	184	3.4	19723	15	US-10-084-846A-5	Sequence 5, Appli
38	183	3.4	306	12	US-10-425-114-68632	Sequence 68632, A
39	183	3.4	332	12	US-10-425-114-53730	Sequence 53730, A
40	183	3.4	1626	14	US-10-202-167-2	Sequence 2, Appli
41	182.5	3.4	217	12	US-10-424-599-172208	Sequence 172208,
42	182.5	3.4	1684	12	US-10-276-774-2398	Sequence 2398, Ap
43	182	3.4	1430	16	US-10-408-765A-1083	Sequence 1083, Ap
44	181.5	3.4	1857	16	US-10-408-765A-2173	Sequence 2173, Ap
45	180.5	3.4	2161	16	US-10-408-765A-1283	Sequence 1283, Ap

#### ALIGNMENTS

##### RESULT 1

US-10-024-368-2  
; Sequence 2, Application US/10024368  
; Publication No. US20030027300A1  
; GENERAL INFORMATION:  
; APPLICANT: THOMPSON, Catherine C.  
; TITLE OF INVENTION: HUMAN HAIRLESS GENE AND PROTEIN  
; FILE REFERENCE: Thompson-20263/0243435  
; CURRENT APPLICATION NUMBER: US/10/024,368  
; CURRENT FILING DATE: 2001-12-21  
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US/09/287,354  
; PRIOR FILING DATE: EARLIER FILING DATE: 1999-04-07  
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US 60/080,888  
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-04-07  
; NUMBER OF SEQ ID NOS: 10  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 2  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-024-368-2

Query Match	100.0%;	Score	5358;	DB	14;	Length	984;
Best Local Similarity	100.0%;	Pred. No.	0;				
Matches	984;	Conservative	0;	Mismatches	0;	Indels	0;
Gaps	0;						
QY	1	FYYKDPISIPRLAKEPLAAAE	PGLFGLNSGGHLORAGEAERPSILHQRD	GEMGAGRQNP	60		
Db	1	FYYKDPISIPRLAKEPLAAAE	PGLFGLNSGGHLORAGEAERPSILHQRD	GEMGAGRQNP	60		
QY	61	LFGLGPDVTVPWTSWACPPGLVHTLGNWAGPGDGNLGYQLGPPATPRCPSPBPVPTQ	120				
Db	61	LFGLGPDVTVPWTSWACPPGLVHTLGNWAGPGDGNLGYQLGPPATPRCPSPBPVPTQ	120				
QY	121	CCSSYPPTKGGDLPGCKQCGLEGASGASPSSEVNKASGPRACPPSHHTKTKTWT	180				
Db	121	CCSSYPPTKGGDLPGCKQCGLEGASGASPSSEVNKASGPRACPPSHHTKTKTWT	180				

QY 601 ERKIQEKALGPGCLRAGPGRKGLGLPLSPVPRPLPPGALLWLQBPQPCRRGHFLFOEH 660  
DB 801 ERKIQEKALGPGCLRAGSGLRGLSLPLSPVTRLSPPGALLWLQBPQPCRRGHFLFOEH 858  
QY 661 WRQGPVLVSGIORTLQGNLWGTALGALGGVQVQALSPGPPQSSLSGTTFWGFSWPE 720  
DB 859 WRQGPVLVSGIORTLQGNLWGTALGALGGVQVQALSPGPPQSSLSGTTFWGFSWPE 918  
QY 721 LRPKSDEGSVLLHRLGDEDTSRVENLAASLPLPEYCALHGKLNLAASLPLGLALRPLE 780  
DB 919 TRPKLDEGSVLLHRLGDEDTSRVENLAASLPLPEYCALHGKLNLAASLPLGLALRPLE 978  
QY 781 POLWAAYGVSHRGLGTKNLCVEVADIVSLVHADTLPDAWHRAQKDFLSLDEGLWS 840  
DB 979 POLWAAYGVSHRGLGTKNLCVEVADIVSLVHADTLPDAWHRAQKDFLSLDEGLWS 1038  
QY 841 PGSQVSTVWHVFRAGDAQRIIRFLQMVCPAGAGALEPGAGSCYLDAGLRRLREMGVS 900  
DB 1039 PGSQVSTVWHVFRAGDAQRIIRFLQMVCPAGAGALEPGAGSCYLDAGLRRLREMGVS 1098  
QY 901 CWTLLQAPGEAVLPAGAPHVQGLVSTVSTVQHFLSPETSALSALCHQGPSLPPDCHL 960  
DB 1099 CWTLLQAPGEAVLPAGAPHVQGLVSTVSTVQHFLSPETSALSALCHQGPSLPPDCHL 1158  
QY 961 LYAQMDWAVFOAVKAVGTLOEAK 984  
DB 1159 LYAQMDWAVFOAVKAVGTLOEAK 1182

RESULT 5  
US-09-287-354-5  
; Sequence 5, Application US/09287354  
; Patent No. 6348348  
; GENERAL INFORMATION:  
; APPLICANT: THOMPSON, Catherine C.  
; TITLE OF INVENTION: HUMAN HAIRLESS GENE AND PROTEIN  
; FILE REFERENCE: Thompson-20263/0243435  
; CURRENT FILING DATE: 1999-04-07  
; EARLIER APPLICATION NUMBER: US 60/080,888  
; NUMBER OF SEQ ID NOS: 10  
; SOFTWARE: Patent In Ver. 2.0  
; SEQ ID NO 5  
; LENGTH: 1207  
; TYPE: PRT  
; ORGANISM: Rat  
US-09-287-354-5

Query Match 75.9%; Score 4067; DB 4; Length 1207;  
Best Local Similarity 77.7%; Pred. No. 0;  
Matches 765; Conservative 65; Mismatches 146; Indels 8; Gaps 5;

QY 1 FYYKDPISPLAKEPLAAAEPLGLNLSGHLQRAGEARPSLHORDGEMGAGROQPCP 60  
DB 232 FYKDPNLRPAKEPLAASEGMLGAGHLQQAADAGSLHORDGETAGROOQPCP 291  
QY 61 LFLGQDTPVMTSPACPPGLVHTLGNWAGFGDGNLGYQLGPPATPRCPSPPEVPTQ 120  
DB 292 VFLGQDTPVMTSPACPPGLVHTLGNWAGFGDGNLGYQLGPPATPRCPSPPEVPTQ 351  
QY 121 CCSSVPPKGGDLGCGKQEGLEGASBPSEVNVKASGPACPSHHTKKTWLT 180  
DB 352 CCSSLPAREGDPGCRKQDPSGSSGSGPESSEERNA-GSRASPPSHHTKKTWLT 410  
QY 181 RHSEQFECPCGPEVEERPVARLRALKRAGSEVQAGMSPAPKPPDPFCTAEOGAG 240  
DB 411 RHSEQFECPCGKGESGATGLRALKRAGSEVQAGMSPAPKPPDPFCTAEOGAG 469  
QY 241 WQEVDTSGNKNVDSDGDEQKQPDQASIQDPLQDIPCLALPAKLACQSCAQAG 300  
DB 470 WQETPSTGSK-AEAQQQEQGRDRIRLRSELEDTSCQHLGAVTQCPSCVQAG 528

QY 301 EGGHACHSQVRRSPGLGELQOQEDTATNSSBEGSGDPDSRLSTGLAKHLISGLGDR 360  
DB 529 EVELTSHSQSHKLPLEKPLEEDSCAT---SEEGGSSPEASINKLAKHLISGLGDR 585  
QY 361 LCRLLRERERALAWAEGGPGPATGSPGIPRCCSCHHGLFNTHWRCPSCSHRLCVAC 420  
DB 586 LCRLLRERERALAWAEGGPGPATGSPGIPRCCSCHHGLFNTHWRCPSCSHRLCVAC 645  
QY 421 GRVAGTGRARERKAGFOESABECTQEAAGHACSLMLTQFVSSQALAEISTAMHVVWVFD 480  
DB 646 GRIAGAKNREKTSRQRQRTDDCAQEAAGHACSLMLTQFVSSQALAEISTAMHVVWVFD 705  
QY 481 IRGHCPOADARVWAPGDAQKQESTOKTPTTPOFSCNGDTHRTKSIKEETPDSEAETAE 540  
DB 706 IRGHCFOVDARVWAPGDAQKQESTOKTPTTPOFSCNGDTHRTKSIKEETPDSEAETAE 765  
QY 541 DRAGRGPLPCPSLCELLASTAVKLCGLHERIHMFAFVPTPALPSDDRTNILDSTIAQVV 600  
DB 766 DRAGRSPLPCPSLCELLASTAVKLCGLHERIHMFAFVPTPALPSDDRTNILDSTIAQVV 825  
QY 601 ERKIQEKALGPGCLRAGPGRKGLGLPLSPVPRPLPPGALLWLQBPQPCRRGHFLFOEH 660  
DB 826 ERKIQEKALGPGCLRAGSGLRGLSLPLSPVTRLSPPGALLWLQBPQPCRRGHFLFOEH 883  
QY 661 WRQGPVLVSGIORTLQGNLWGTALGALGGVQVQALSPGPPQSSLSGTTFWGFSWPE 720  
DB 884 WRQGPVLVSGIORTLQGNLWGTALGALGGVQVQALSPGPPQSSLSGTTFWGFSWPE 943  
QY 721 LRPKSDEGSVLLHRLGDEDTSRVENLAASLPLPEYCALHGKLNLAASLPLGLALRPLE 780  
DB 944 ARPXLDEGSVLLHRLGDEDTSRVENLAASLPLPEYCALHGKLNLAASLPLGLALRPLE 1003  
QY 781 POLWAAYGVSHRGLGTKNLCVEVADIVSLVHADTLPDAWHRAQKDFLSLDEGLWS 840  
DB 1004 POLWAAYGVSHRGLGTKNLCVEVADIVSLVHADTLPDAWHRAQKDFLSLDEGLWS 1063  
QY 841 PGSQVSTVWHVFRAGDAQRIIRFLQMVCPAGAGALEPGAGSCYLDAGLRRLREMGVS 900  
DB 1064 PGSQVSTVWHVFRAGDAQRIIRFLQMVCPAGAGALEPGAGSCYLDAGLRRLREMGVS 1123  
QY 901 CWTLLQAPGEAVLPAGAPHVQGLVSTVSTVQHFLSPETSALSALCHQGPSLPPDCHL 960  
DB 1124 CWTLLQAPGEAVLPAGAPHVQGLVSTVSTVQHFLSPETSALSALCHQGPSLPPDCHL 1183  
QY 961 LYAQMDWAVFOAVKAVGTLOEAK 984  
DB 1184 LYAQMDWAVFOAVKAVGTLOEAK 1207

RESULT 6  
US-09-252-991A-22176  
; Sequence 22176, Application US/09252991A  
; Patent No. 6551795  
; GENERAL INFORMATION:  
; APPLICANT: Marc J. Rubenfield et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
; FILE REFERENCE: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS  
; FILE REFERENCE: 107196.136  
; CURRENT APPLICATION NUMBER: US/09/252,991A  
; PRIORITY FILING DATE: 1999-02-18  
; PRIOR APPLICATION NUMBER: US 60/074,788  
; PRIOR FILING DATE: 1998-02-18  
; PRIOR APPLICATION NUMBER: US 60/094,190  
; PRIOR FILING DATE: 1998-07-27  
; NUMBER OF SEQ ID NOS: 33142  
; SEQ ID NO 22176  
; LENGTH: 705  
; TYPE: PRT  
; ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-22176

Query Match 4.1%; Score 219.5; DB 4; Length 705;



EARLIER APPLICATION NUMBER: US 60/080,888  
EARLIER FILING DATE: 1998-04-07  
NUMBER OF SEQ ID NOS: 10  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 3  
LENGTH: 1189  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-287-354-3

Query Match 98.7%; Score 5290; DB 4; Length 1189;  
Best Local Similarity 99.0%; Pred. No. 0;  
Matches 974; Conservative 1; Mismatches 9; Indels 0; Gaps 0;

QY 1 FYYKDSIPRLAKEPLAAAEPLGLNSGHLORAGEAERPSLHQRDGEAGRQONPCP 60  
Db FYYKDSIPRLAKEPLAAAEPLGLNSGHLORAGEAERPSLHQRDGEAGRQONPCP 265  
QY 61 LFLGQDPTVPWTSWAPCPGLVHTLGNVWAGPDGMLGYQLGPPATPRCPSPPEPVTTQ 120  
Db LFLGQDPTVPWTSWAPCPGLVHTLGNVWAGPDGMLGYQLGPPATPRCPSPPEPVTTQ 325  
QY 121 CCSSYPPTKGGDLGPCGKQCEGLGEGGASGASEPSEEVNKASGPRACPPSHHTKLKKTWLT 180  
Db CCSSYPPTKGGDLGPCGKQCEGLGEGGASGASEPSEEVNKASGPRACPPSHHTKLKKTWLT 385  
QY 181 RHSEQECPRGCPVEERPVARLRALKRAGSPVQAGMSPAPKRPDPPTGTAEOAGG 240  
Db RHSEQECPRGCPVEERPVARLRALKRAGSPVQAGMSPAPKRPDPPTGTAEOAGG 445  
QY 241 WQVTRTSIGNKDVDSGQHDQKQPGDQASLQDPGLQDIPCLALPAKLAQCQSCAAAG 300  
Db LQEVTRTSIGNKDVDSGQHDQKQPGDQASLQDPGLQDIPCLALPAKLAQCQSCAAAG 505  
QY 301 EGGGHACHSQVRRSPPLGGELQOEEDTATNSSSEEGSGSPDSTGLAKHLLSGLD 360  
Db EGGGHACHSQVRRSPPLGGELQOEEDTATNSSSEEGSGSPDSTGLAKHLLSGLD 565  
QY 361 LCELLRERREALAWAREGQGPATVGTSGIPRCCSRCHGLFNTWRCPCRSRHLVCAC 420  
Db LCELLRERREALAWAREGQGPATVGTSGIPRCCSRCHGLFNTWRCPCRSRHLVCAC 625  
QY 421 GRVAGTGRAREKAGFOEQSAEECTQAGHAACSLMTQFVSSQALAEELSTAMHVQVKFD 480  
Db GRVAGTGRAREKAGFOEQSAEECTQAGHAACSLMTQFVSSQALAEELSTAMHVQVKFD 685  
QY 481 IRGHCPQADARWAPGDAGQKQESTQKTPPTQPCNGDTHRTKSIKEETPDSETPAE 540  
Db IRGHCPQADARWAPGDAGQKQESTQKTPPTQPCNGDTHRTKSIKEETPDSETPAE 745  
QY 541 DRAGRPLPCPSICELLASTAVKLCIGHRIHMAFAPVTPALPDSDDRTNILDIIAQVV 600  
Db DRAGRPLPCPSICELLASTAVKLCIGHRIHMAFAPVTPALPDSDDRTNILDIIAQVV 805  
QY 601 ERKIOBKALGPGLRAGPGRKGLGLPLSPVRPLPPLPGALLMLQBPQCPRRGFHLFOEH 660  
Db ERKIOBKALGPGLRAGPGRKGLGLPLSPVRPLPPLPGALLMLQBPQCPRRGFHLFOEH 865  
QY 661 WRQGPVLVSGIORTLQGNLWGTALGALGQVQALSPGLPPQSSIGSTTFWEGFSWPE 720  
Db WRQGPVLVSGIORTLQGNLWGTALGALGQVQALSPGLPPQSSIGSTTFWEGFSWPE 925  
QY 721 LRPKSDEGSVLILHRLGALGDETSRVENLAASLPLPYCALHGKLNLASVLPGLALRPLE 780  
Db LRPKSDEGSVLILHRAFGDETSRVENLAASLPLPYCALHGKLNLASVLPGLALRPLE 985  
QY 781 POLMAAYGVSPIRGHIGTKNLCEVEADLVSLVHARTPLPAMHRAQKDFLSGLDGLGWS 840  
Db POLMAAYGVSPIRGHIGTKNLCEVEADLVSLVHARTPLPAMHRAQKDFLSGLDGLGWS 1045  
QY 841 PGSQVSTVHVHFAQAQARRRFLQWVCAGALEPGAGSCYLDAGLRRLRREEWGVS 900  
Db PGSQVSTVHVHFAQAQARRRFLQWVCAGALEPGAGSCYLDAGLRRLRREEWGVS 1105

QY 901 CWTLLQAPGEAVLVPAGAPHQVQGLVSTVSTQHFILSPETSALSQAQLCHQSPSLPPDCHL 960  
Db CWTLLQAPGEAVLVPAGAPHQVQGLVSTVSTQHFILSPETSALSQAQLCHQSPSLPPDCHL 1165  
QY 961 LYAQMDWAVFQAVKVAVTGLQEA 984  
Db LYAQMDWAVFQAVKVAVTGLQEA 1189

## RESULT 4

US-09-287-354-6  
Sequence 6, Application US/09287354  
Patent No. 6348348  
GENERAL INFORMATION:  
APPLICANT: THOMPSON, Catherine C.  
TITLE OF INVENTION: HUMAN HAIRLESS GENE AND PROTEIN  
FILE REFERENCE: Thompson-20263/0243435  
CURRENT APPLICATION NUMBER: US/09/287,354  
CURRENT FILING DATE: 1999-04-07  
EARLIER APPLICATION NUMBER: US 60/080,888  
EARLIER FILING DATE: 1998-04-07  
NUMBER OF SEQ ID NOS: 10  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 6  
LENGTH: 1182  
TYPE: PRT  
ORGANISM: Mouse  
US-09-287-354-6

Query Match 76.6%; Score 4102.5; DB 4; Length 1182;  
Best Local Similarity 78.4%; Pred. No. 0;  
Matches 771; Conservative 62; Mismatches 144; Indels 7; Gaps 5;

QY 1 FYYKDSIPRLAKEPLAAAEPLGLNSGHLORAGEAERPSLHQRDGEAGRQONPCP 60  
Db FYYKDSIPRLAKEPLAAAEPLGLNSGHLORAGEAERPSLHQRDGEAGRQONPCP 263  
QY 61 LFLGQDPTVPWTSWAPCPGLVHTLGNVWAGPDGMLGYQLGPPATPRCPSPPEPVTTQ 120  
Db VFLGYDPTVPWTSWAPCPGLVHTLGNVWAGPDGMLGYQLGPPATPRCPSPPEPVTTQ 323  
QY 121 CCSSYPPTKGGDLGPCGKQCEGLGEGGASGASEPSEEVNKASGPRACPPSHHTKLKKTWLT 180  
Db CCSSHLPAREGLGPCRQKQDSPEGGSGSGSESEERNKADS-RACPPSHHTKLKKTWLT 382  
QY 181 RHSEQECPRGCPVEERPVARLRALKRAGSPVQAGMSPAPKRPDPPTGTAEOAGG 240  
Db RHSEQECPRGCPVEERPVARLRALKRAGSPVQAGMSPAPKRPDPPTGTAEOAGG 442  
QY 241 WQVTRTSIGNKDVDSGQHDQKQPGDQASLQDPGLQDIPCLALPAKLAQCQSCAAAG 300  
Db WQETPETIIGSK-ABAEQOEQGRPRDGRIRLQESRLVDTSCQHLHAGVTQCQSCVQAAG 501  
QY 301 EGGGHACHSQVRRSPPLGGELQOEEDTATNSSSEEGSGSPDSTGLAKHLLSGLD 360  
Db EVGVLGTGHSQKSRSP-EEKQLEEDSSATSEEGGGPGPEASLNKLAHLLSGLD 560  
QY 361 LCELLRERREALAWAREGQGPATVGTSGIPRCCSRCHGLFNTWRCPCRSRHLVCAC 420  
Db LCELLRERREALAWAREGQGPATVGTSGIPRCCSRCHGLFNTWRCPCRSRHLVCAC 620  
QY 421 GRVAGTGRAREKAGFOEQSAEECTQAGHAACSLMTQFVSSQALAEELSTAMHVQVKFD 480  
Db GRVAGTGRAREKAGFOEQSAEECTQAGHAACSLMTQFVSSQALAEELSTAMHVQVKFD 680  
QY 481 IRGHCPQADARWAPGDAGQKQESTQKTPPTQPCNGDTHRTKSIKEETPDSETPAE 540  
Db IRGHCPQADARWAPGDAGQKQESTQKTPPTQPCNGDTHRTKSIKEETPDSETPAE 740  
QY 541 DRAGRPLPCPSICELLASTAVKLCIGHRIHMAFAPVTPALPDSDDRTNILDIIAQVV 600  
Db DRAGRPLPCPSICELLASTAVKLCIGHRIHMAFAPVTPALPDSDDRTNILDIIAQVV 800



Db 361 LCRLLRRERREALAWAQREGGQPAVGTGSDGPGIPLCCSCHHGLFNTHWRCPRCSHRLCVAC 420  
QY 421 GRVAGTGRAREKAGFQESAEECTQEAAGAACSLMLTQFVSSQALAEIATAMHQQVWKFD 480  
Db 421 GRVAGTGRAREKAGFQESAEECTQEAAGAACSLMLTQFVSSQALAEIATAMHQQVWKFD 480  
QY 481 IRGHCPQADARWAPGDAGQOKESTQKTPPTPOPCNGDTHRTKSIKEETPDSEATPAE 540  
Db 481 IRGHCPQADARWAPGDAGQOKESTQKTPPTPOPCNGDTHRTKSIKEETPDSEATPAE 540  
QY 541 DRAGRGLPCPSLCELLASTAVKLCIHERIHMFAFVTPALPDDRTITNLSIIAQVV 600  
Db 541 DRAGRGLPCPSLCELLASTAVKLCIHERIHMFAFVTPALPDDRTITNLSIIAQVV 600  
QY 601 ERKIQEALGPGLRAGPGLRKGLGLPLSPVPRPLPPGALLWLQEPQCPRRGFHLFQEH 660  
Db 601 ERKIQEALGPGLRAGPGLRKGLGLPLSPVPRPLPPGALLWLQEPQCPRRGFHLFQEH 660  
QY 661 WRQOPVLVSGIORTLOGLWGTALGALGQVQALSPGPPQPSLGSSTTFWEGFSWPE 720  
Db 661 WRQOPVLVSGIORTLOGLWGTALGALGQVQALSPGPPQPSLGSSTTFWEGFSWPE 720  
QY 721 LRPKSDGVSLLHRLAGDEDTSRVENLAASLPLPEYCALHGKLNLSYLPGLALRPLE 780  
Db 721 LRPKSDGVSLLHRLAGDEDTSRVENLAASLPLPEYCALHGKLNLSYLPGLALRPLE 780  
QY 781 POLWAAAYGVSPHRLGKTKNLCVEVADIVSLVHADTPLPAWHRAQKDFLSGLDGEGLWS 840  
Db 781 POLWAAAYGVSPHRLGKTKNLCVEVADIVSLVHADTPLPAWHRAQKDFLSGLDGEGLWS 840  
QY 841 PGSOVSTVWHVFRADQORIRRFLOMVCPCAGALEPGAGSCYLDAGLRRLREEWGVS 900  
Db 841 PGSOVSTVWHVFRADQORIRRFLOMVCPCAGALEPGAGSCYLDAGLRRLREEWGVS 900  
QY 901 CWTLLQAPGEAVLPAGAPHQVQGLVSTVSTQHFSLPETSALSQALCHQGPSLPPDCHL 960  
Db 901 CWTLLQAPGEAVLPAGAPHQVQGLVSTVSTQHFSLPETSALSQALCHQGPSLPPDCHL 960  
QY 961 LYAQMNAVFOAVKAVGTTLQEAK 984  
Db 961 LYAQMNAVFOAVKAVGTTLQEAK 984

## RESULT 2

US-09-287-354-4  
; Sequence 4, Application US/09287354  
; Patent No. 6348348  
; GENERAL INFORMATION:  
; APPLICANT: THOMPSON, Catherine C.  
; TITLE OF INVENTION: HUMAN HAIRLESS GENE AND PROTEIN  
; FILE REFERENCE: Thompson-20263/0243435  
; CURRENT APPLICATION NUMBER: US/09/287,354  
; EARLIER FILING DATE: 1999-04-07  
; EARLIER APPLICATION NUMBER: US 60/080,888  
; EARLIER FILING DATE: 1998-04-07  
; NUMBER OF SEQ ID NOS: 10  
; SOFTWARE: Patent in ver. 2.0  
; SEQ ID NO 4  
; LENGTH: 1189  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-287-354-4

Query Match 99.7%; Score 5343; DB 4; Length 1189;  
Best Local Similarity 99.8%; Pred. No. 0;  
Matches 982; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 FYYKDPSTPRIAKEPLAAEPGLFGINSGGHLOAGAEERPSLHORDGEMGAGROQNP 60  
Db 206 FYYKDPSTPRIAKEPLAAEPGLFGINSGGHLOAGAEERPSLHORDGEMGAGROQNP 265  
QY 61 LFLGQPDTPVWTSWPAWPGDGLVHTLGNVWAGPGDGLVSTVSTQHFSLPETSALSQALCHQGPSLPPDCHL 120

Db 266 LFLGQPDTPVWTSWPAWPGDGLVHTLGNVWAGPGDGLVSTVSTQHFSLPETSALSQALCHQGPSLPPDCHL 325  
QY 121 CCSSYPPTKGGDLGPGCKQEGLEGASGASEPSEEVNKASGPRACPPSHHTLKTWTLT 180  
Db 326 CCSSYPPTKGGDLGPGCKQEGLEGASGASEPSEEVNKASGPRACPPSHHTLKTWTLT 385  
QY 181 RHSEQFPCPCPEVERPVARLALRKRAGSPVQGMGSPAPKRPDPDPFGTAEOQAGG 240  
Db 386 RHSEQFPCPCPEVERPVARLALRKRAGSPVQGMGSPAPKRPDPDPFGTAEOQAGG 445  
QY 241 WQEVDRDSIGNKVDVSGQHDQKGPQDQASLODIPCLALPAKLPAKQCSQAQAG 300  
Db 446 WQEVDRDSIGNKVDVSGQHDQKGPQDQASLODIPCLALPAKLPAKQCSQAQAG 505  
QY 301 EGGHACHSQOVRPSPLGGBELQBEEDTATNSSSEEGPGSDSLSTGLAKHLSGLGDR 360  
Db 506 EGGHACHSQOVRPSPLGGBELQBEEDTATNSSSEEGPGSDSLSTGLAKHLSGLGDR 565  
QY 361 LCRLLRRERREALAWAQREGGQPAVGTGSDGPGIPLCCSCHHGLFNTHWRCPRCSHRLCVAC 420  
Db 566 LCRLLRRERREALAWAQREGGQPAVGTGSDGPGIPLCCSCHHGLFNTHWRCPRCSHRLCVAC 625  
QY 421 GRVAGTGRAREKAGFQESAEECTQEAAGAACSLMLTQFVSSQALAEIATAMHQQVWKFD 480  
Db 626 GRVAGTGRAREKAGFQESAEECTQEAAGAACSLMLTQFVSSQALAEIATAMHQQVWKFD 685  
QY 481 IRGHCPQADARWAPGDAGQOKESTQKTPPTPOPCNGDTHRTKSIKEETPDSEATPAE 540  
Db 686 IRGHCPQADARWAPGDAGQOKESTQKTPPTPOPCNGDTHRTKSIKEETPDSEATPAE 745  
QY 541 DRAGRGLPCPSLCELLASTAVKLCIHERIHMFAFVTPALPDDRTITNLSIIAQVV 600  
Db 746 DRAGRGLPCPSLCELLASTAVKLCIHERIHMFAFVTPALPDDRTITNLSIIAQVV 805  
QY 601 ERKIQEALGPGLRAGPGLRKGLGLPLSPVPRPLPPGALLWLQEPQCPRRGFHLFQEH 660  
Db 806 ERKIQEALGPGLRAGPGLRKGLGLPLSPVPRPLPPGALLWLQEPQCPRRGFHLFQEH 865  
QY 661 WRQOPVLVSGIORTLOGLWGTALGALGQVQALSPGPPQPSLGSSTTFWEGFSWPE 720  
Db 866 WRQOPVLVSGIORTLOGLWGTALGALGQVQALSPGPPQPSLGSSTTFWEGFSWPE 925  
QY 721 LRPKSDGVSLLHRLAGDEDTSRVENLAASLPLPEYCALHGKLNLSYLPGLALRPLE 780  
Db 926 LRPKSDGVSLLHRLAGDEDTSRVENLAASLPLPEYCALHGKLNLSYLPGLALRPLE 985  
QY 781 POLWAAAYGVSPHRLGKTKNLCVEVADIVSLVHADTPLPAWHRAQKDFLSGLDGEGLWS 840  
Db 986 POLWAAAYGVSPHRLGKTKNLCVEVADIVSLVHADTPLPAWHRAQKDFLSGLDGEGLWS 1045  
QY 841 PGSOVSTVWHVFRADQORIRRFLOMVCPCAGALEPGAGSCYLDAGLRRLREEWGVS 900  
Db 1046 PGSOVSTVWHVFRADQORIRRFLOMVCPCAGALEPGAGSCYLDAGLRRLREEWGVS 1105  
QY 901 CWTLLQAPGEAVLPAGAPHQVQGLVSTVSTQHFSLPETSALSQALCHQGPSLPPDCHL 960  
Db 1106 CWTLLQAPGEAVLPAGAPHQVQGLVSTVSTQHFSLPETSALSQALCHQGPSLPPDCHL 1165  
QY 961 LYAQMNAVFOAVKAVGTTLQEAK 984  
Db 1166 LYAQMNAVFOAVKAVGTTLQEAK 1189

## RESULT 3

US-09-287-354-3  
; Sequence 3, Application US/09287354  
; Patent No. 6348348  
; GENERAL INFORMATION:  
; APPLICANT: THOMPSON, Catherine C.  
; TITLE OF INVENTION: HUMAN HAIRLESS GENE AND PROTEIN  
; FILE REFERENCE: Thompson-20263/0243435  
; CURRENT APPLICATION NUMBER: US/09/287,354  
; CURRENT FILING DATE: 1999-04-07

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 29, 2004, 10:25:04 ; Search time 23 Seconds  
(without alignments)  
2208.694 Million cell updates/sec

Title: US-10-024-368-2

Perfect score: 5358

Sequence: 1 FYYKPSIPRLAKEPLAAAE.....MDWAVFQAVKVAVGLQEAK 984

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents AA: \*  
1: /cgn2\_6/prodata/2/1aa/5A COMB.pep: \*  
2: /cgn2\_6/prodata/2/1aa/5B COMB.pep: \*  
3: /cgn2\_6/prodata/2/1aa/6A COMB.pep: \*  
4: /cgn2\_6/prodata/2/1aa/6B COMB.pep: \*  
5: /cgn2\_6/prodata/2/1aa/6C COMB.pep: \*  
6: /cgn2\_6/prodata/2/1aa/6D COMB.pep: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	5358	100.0	984	US-09-287-354-2	Sequence 2, Appli
2	5343	99.7	1189	US-09-287-354-4	Sequence 4, Appli
3	5290	98.7	1189	US-09-287-354-3	Sequence 3, Appli
4	4102.5	76.6	1182	US-09-287-354-6	Sequence 6, Appli
5	4067	75.9	1207	US-09-287-354-5	Sequence 5, Appli
6	219.5	4.1	705	US-09-252-991A-22176	Sequence 22176, A
7	184.5	3.4	1384	US-08-976-255-11	Sequence 11, Appl
8	183	3.4	2294	US-09-252-991A-17231	Sequence 17231, A
9	179	3.3	605	US-09-252-991A-22793	Sequence 22793, A
10	179	3.3	917	US-09-252-991A-25101	Sequence 25101, A
11	176.5	3.3	1043	US-09-252-991A-28885	Sequence 28885, A
12	176.5	3.3	1706	US-09-252-991A-31760	Sequence 31760, A
13	172.5	3.2	841	US-09-252-991A-36919	Sequence 36919, A
14	171.5	3.2	1149	US-09-252-991A-45557	Sequence 45557, A
15	170	3.2	1031	US-09-252-991A-18365	Sequence 18365, A
16	168	3.1	553	US-09-252-991A-32621	Sequence 32621, A
17	168	3.1	780	US-09-252-991A-18846	Sequence 18846, A
18	168	3.1	1317	US-09-083-521-7	Sequence 7, Appli
19	167	3.1	822	US-09-252-991A-22479	Sequence 22479, A
20	166.5	3.1	1002	US-09-252-991A-27980	Sequence 27980, A
21	166	3.1	1366	US-09-585-887-10	Sequence 10, Appl
22	166	3.1	1366	US-09-289-578-10	Sequence 10, Appl
23	166	3.1	1694	US-08-494-168-2	Sequence 2, Appli
24	164	3.1	633	US-09-252-991A-23497	Sequence 23497, A
25	163	3.0	955	US-09-252-991A-24254	Sequence 24254, A
26	163	3.0	1358	US-07-945-283-2	Sequence 2, Appli
27	162	3.0	916	US-09-252-991A-31242	Sequence 31242, A

28	162	3.0	1088	3	US-09-130-242-2	Sequence 2, Appli
29	162	3.0	1366	3	US-08-963-825-19	Sequence 19, Appl
30	162	3.0	1366	4	US-09-500-811-19	Sequence 19, Appl
31	162	3.0	1366	4	US-09-570-573-19	Sequence 19, Appl
32	162	3.0	1366	4	US-09-548-608-19	Sequence 19, Appl
33	161.5	3.0	977	4	US-09-252-991A-16655	Sequence 16655, A
34	160.5	3.0	1057	3	US-08-931-820-4	Sequence 4, Appli
35	160	3.0	545	4	US-09-252-991A-30417	Sequence 30417, A
36	159.5	3.0	957	4	US-09-252-991A-20408	Sequence 20408, A
37	159	3.0	474	4	US-09-252-991A-28084	Sequence 28084, A
38	159	3.0	743	4	US-09-252-991A-28327	Sequence 28327, A
39	159	3.0	1180	4	US-09-252-991A-32464	Sequence 32464, A
40	159	3.0	1418	3	US-08-963-825-20	Sequence 20, Appl
41	159	3.0	1418	3	US-09-010-999-1	Sequence 1, Appli
42	159	3.0	1418	4	US-09-500-811-20	Sequence 20, Appl
43	159	3.0	1418	4	US-09-570-573-20	Sequence 20, Appl
44	159	3.0	1418	4	US-09-548-608-20	Sequence 20, Appl
45	158.5	3.0	488	4	US-09-252-991A-24759	Sequence 24759, A

#### ALIGNMENTS

RESULT 1  
US-09-287-354-2  
; Sequence 2, Application US/09287354  
; Patent No. 6348348  
; GENERAL INFORMATION:  
; APPLICANT: THOMPSON, Catherine C.  
; TITLE OF INVENTION: HUMAN HAIRLESS GENE AND PROTEIN  
; FILE REFERENCE: Thompson-20263/0243435  
; CURRENT APPLICATION NUMBER: US/09/287,354  
; EARLIER FILING DATE: 1999-04-07  
; EARLIER APPLICATION NUMBER: US 60/080,888  
; EARLIER FILING DATE: 1998-04-07  
; NUMBER OF SEQ ID NOS: 10  
; SOFTWARE: Patent in Ver. 2.0  
; SEQ ID NO 2  
; LENGTH: 984  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-287-354-2

Query Match		100.0%;	Score 5358;	DB 4;	Length 984;
Best Local Similarity		100.0%;	Pred. No. 0;		
Matches 984;		Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1	FYYKPSIPRLAKEPLAAAE	PGLFGLNSGGHLQRA	GEAERPSLHQRD	GEMGAGRQNP
DB	1	FYYKPSIPRLAKEPLAAAE	PGLFGLNSGGHLQRA	GEAERPSLHQRD	GEMGAGRQNP
QY	61	LFLGQPDTPVWTSW	PACPPGLVHTLGNVW	AGDGNLGYQLGPP	ATPCPSPEPPVT
DB	61	LFLGQPDTPVWTSW	PACPPGLVHTLGNVW	AGDGNLGYQLGPP	ATPCPSPEPPVT
QY	121	CCSSYPPTKGGDLG	PCGKQCEGLGAS	GPSEEVNKASGP	RACPSSHHTLKT
DB	121	CCSSYPPTKGGDLG	PCGKQCEGLGAS	GPSEEVNKASGP	RACPSSHHTLKT
QY	181	RHSEFCEPRGCP	VEEVRVAILRAL	KRAGSEVQGM	SGSPAPKRPDP
DB	181	RHSEFCEPRGCP	VEEVRVAILRAL	KRAGSEVQGM	SGSPAPKRPDP
QY	241	WQEVRTDTIGNK	VDVSGQHD	EOKGPDGQAS	IQDPLQDIPCLAL
DB	241	WQEVRTDTIGNK	VDVSGQHD	EOKGPDGQAS	IQDPLQDIPCLAL
QY	301	EGGHACHSQVRR	SPLGGLGELQ	EEEDATNS	SEEGSGSPDS
DB	301	EGGHACHSQVRR	SPLGGLGELQ	EEEDATNS	SEEGSGSPDS
QY	361	LCRLRRREALA	WQREGQ	PAVTDSPG	IPRCCSRCHHGL
DB	361	LCRLRRREALA	WQREGQ	PAVTDSPG	IPRCCSRCHHGL

324	CCSSYPPIKDGGLGPCCKCEGLDGVASGASEEYNKASGRACPPGHHTKLKKTWLT	383
181	RHSEQEPCPGCEVEERPVARLRAKRAAGSPVQAGMGSPAKRPDPDPFGTAEOGAGG	240
384	RHSEQEPCPGCEFAEERPVQALRAKRAAGSPELQAVGGPAPKRPDPDPFGTAEOGAGG	443
241	WQEVRTSITGNKOVDSQHDQKGPQDQASLQDPGLQDIPCIALPAKLACQCSQAQAG	300
444	WQEVRTSITGNKEADSQHDQRPDQASLQDPGLQEIPIALPAKLACQCSQAQAG	503
301	EGGGHACHSOQVRSPLGGELQOEEDATNSSBEGSGSPDSELSTGLIAKHLLSLGDR	360
504	EGGGPAGHFQOVRRSPUGGELQOEEDTAANSSBEGSGPDGRLSTGLSKHLLSSLGDR	563
361	LCRLLRERREALAWAQREGQGPVTDSPGIPRCCSRCHHGLFNTHWRCPRCSHRLCVAC	420
564	LCRLLRERREALAWAQREGQGPVTDENPGIPRCCSRCHHGLFNTHWRCPRCSHRLCVAC	623
421	GRVAGTRAREKAGFQBSAECTQEBAGHAACSLMLTFQVSSQALABELSTAMHQVWKFD	480
624	GRVAGTRAREKAGSRQEETCTQEBAGHAACSLTLTFQVSSQALABELSTAMHQVWKFD	683
481	IRGHCPQADARVWAPGDAGQCKESTOKTPTTPQPSNGDTHRTKSTKETPDSAEPTAE	540
684	IRGHCPQADARVWAPGDAGQCKESTOKTPTTPQPSNGDTHRTKSTKETPDSAEPTAE	743
541	DRAGRGPLPCPSLCELLASTAVKLCLGHERIHMAFAPVTPALPSDDRIITNLDSIIAQVV	600
744	DRAGQAPLPCPSLCELLASTAVKLCLGHERIHMAFAPVTPALPSDDRIITNLDSIIAQVV	803
601	ERKIQEKALQPLRAGPGLRKGGLPLSPVVRPLPPGALLWTQEPQPCPRRGPHLFOEH	660
804	ERKIQEKALQPLRAGPGLRKGGLPLSPVVRPLPPGALLWTQEPQPRPQRFHLPFOEH	863
661	WRQGPVLVSGIORTLOGNLWGTEALGAGQVQALSPLGPPQSSLSGTSFTWEGFSWPE	720
864	WRQGPVLVSGIORTLOGNLWGTEALGAGQVQALSPLGPPQSSLSGTSFTWEGFSWPE	923
721	LRPKSDEGSVLLHRLAGDEDTSRVENLAASLPLPEYCALHGKLNLSYLPPLGLALRPLE	780
924	LRPKSDEGSVLLHRLAGDEDTSRVENLAASLPLPEYCAHRGKLNLSYLPPLGLALRPLE	983
781	PQIWAAYGVSPHGRHLGTKNLCVEADVSLIVHADTPLPAWHRAQKDPFSLGDLDEGLWS	840
984	PQIWAAYGVSPHGRHLGTKNLCVEADVSLIVHADTPLPAWHRAQKDPFSLGDLDEGLWS	1043
841	PGSOVSTVHVHFAQDAQRIRRFLOWMCPAGAGALEPGASCYLDAGLRRLREENGVS	900
1044	PGSOVSTVHVHFAQDAQRIRRFLOWMCPAGAGALEPGASCYLDAGLRRLREENGVS	1103
901	CWTLIQAPGEAVLPVAPGAPHQVQGLVSTVSTQHFLSPETSAISAQCHQOGPSLPDPCHL	960
1104	CWTLIQAPGEAVLPVAPGAPHQVQGLVSTVSTQHFLSPETSAISAQCHQOGPSLPDPCHL	1163
961	LYAQMDWAVFQAVKVAVGTIQAOK	984
1164	LYAQMDWAVFQAVKVAVGTIQAOK	1187

## RESULT 2

Q8HZ76	PRELIMINARY	PRT; 709 AA.
ID	Q8HZ76	
AC	Q8HZ76	
DT	01-MAR-2003 (TREMBlrel. 23, Created)	
DT	01-MAR-2003 (TREMBlrel. 23, Last sequence update)	
DT	01-OCT-2003 (TREMBlrel. 25, Last annotation update)	
DE	Hairless (Fragment).	
DE	Ovis aries (Sheep).	
OS	Ovis aries (Sheep).	
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Eutelesostomi;	
OC	Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoides;	
OC	Bovidae; Caprinae; Ovis.	
QX	NCBI TaxID=9940;	

RN	[1]	SEQUENCE FROM N.A.
RP		Finochiaro R., Portolano B., Damiani G., Caroli A., Budelli E.,
RA		Bolla P., Pagnacco G.;
RA		"The hairless (hr) gene is involved in the congenital hypotrichosis of
RT		Valle del Belice sheep.";
RL		Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
DR		EMBL; AY130969; AA05753.1; -.
DR		GO; GO:0005489; F:electron transporter activity; IEA.
DR		GO; GO:0006118; P:electron transport; IEA.
DR		InterPro; IPR000345; CyC_heme_BS.
DR		PROSITE; PS00190; CYTOCHROME_C; 1.
FT		NON TER 709
SQ		SEQUENCE 709 AA; 75435 MW; 838EC14E3A1B0071 CRC64;
Query Match                  34.3%; Score 1837; DB 6; Length 709;		
Best Local Similarity      68.2%; Pred.No.3.le-120;		
Matches 34; Conservative    32; Mismatches 124; Indels     6; Gaps       3		
Qy	1	FYKDPSPIRLAKEPLAAAPGLFGLNSGGHLQRAGEAPRSLHORDGEMAGROONPCP 60
Dd	206	FYHKDPSILRLANDPLATVEPLLGSAPRHQLQTGEVHEHTLHQDGETGVGRHNLCPP 265
Qy	61	LFTGQPTVTWTSWPACPGVLVHTLGNWAGPDGNTLGYLGPATPRCPPEPVTVQG 120
Dd	266	LLLGHPTDVARTPWTCPSGLVHTLGNWAVPGGSLGYQLGPSAATRCPSGPPPTQAG 325
Qy	121	CCSSYPPTKGDLGPCCKCOEGLEGASASEPSEEVNKASGRACPPSHHTKLKKTWLT 180
Dd	326	YCSSHPPARD---DPCQQCFEDLEGATSEPSESSEAAYKTSPRACLPSHTKLKKTWLT 382
Qy	181	RHSQEPFCRGCEVERPERVARLRALKRAGSPEVOGAMGSPAKRPDPFPCTAEQGAGG 240
Dd	383	RHSQFCPDSCEGESPAQLPARRRSSFEVGTASSFAAKRTPGFPGSVCGGARG 442
Qy	241	WEYRDITSIGNKDVSQHDCEKQPQDQSALQDPGLQDIPLCLAPAKLAQCSCQAQAG 300
Dd	443	RQEVLDSVFNGK-AETQRDDHRCDRGDDGASLQAPGDITGVPPVAGITCCGCAQAG 501
Qy	301	EGGHACHISQOVRSPIGELIQBEDPATNSSSESGPSGDSRLSTGLAKHLISGLGR 360
Dd	502	ETGPGAHYSQPLSRLLLGGEQPOEEDSA--ASSKEGGGSGPEAGLSVGLAKHLISALGR 559
Qy	361	LCRLIRERREALWAQREGOGPATVGDPGTPRCCSRCHHGLFNTHWCPCRSHELVCAC 420
Dd	560	LCRVLRERREALVMAHQGAQPVGEDDPSLPCCSCHHGLFNTHWCPCRSHELVCVC 619
Qy	421	GRVAGTGRAREKAGFOQSAECTQEAAGHAACSMLTFQVSSQALAEIESTAMHVQWKFD 480
Dd	620	GRMAGASTREKAGSQOSTECKQEQAGHSGSLMTQFVSSQPLAEISAAMHVQWKFD 679
Qy	481	IRGHCPQADARVWAPGDAGQOKBESTQKTP 510
Dd	680	IRGHCPQADARVWTAGDGLLOKPEKTP 709

```

RESULT 3
Q863E6
ID Q863E6 PRELIMINARY; PRT; 342 AA.
AC Q863E6;
DT 01-JUN-2003 (TRENBLrel. 24, Created)
DT 01-JUN-2003 (TRENBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)
DE Hairless (Fragment).
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OC NCBI_TaxID=9823;
RN [1]
RP SEQUENCE FROM N.A.
RF Fernandez A.I., Silio L., Rodriguez C., Noguera J.L., Sanchez A.,
RA Ovilo C.;
RT "Hairless gene porcine, mRNA (exons 11 to 19).";

```

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 29, 2004, 10:23:29 ; Search time 50 Seconds  
(without alignments)  
6209.398 Million cell updates/sec

Title: US-10-024-368-2  
Perfect score: 5358  
Sequence: 1 FYIKDPSIPRLAKEPLAAAE.....MDWAVFQAVKVAVGLTQEAQ 984

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

- SPTREMBL.25.\*
- 1: sp\_archaea.\*
  - 2: sp\_bacteria.\*
  - 3: sp\_fungi.\*
  - 4: sp\_human.\*
  - 5: sp\_invertebrate.\*
  - 6: sp\_mammal.\*
  - 7: sp\_mhc.\*
  - 8: sp\_organelle.\*
  - 9: sp\_phase.\*
  - 10: sp\_plant.\*
  - 11: sp\_rodent.\*
  - 12: sp\_virus.\*
  - 13: sp\_vertebrate.\*
  - 14: sp\_unclassified.\*
  - 15: sp\_rvirus.\*
  - 16: sp\_bacteriap.\*
  - 17: sp\_archaeap.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	5119	95.5	1187	6 Q8WNL9	Q8wnl9 macaca mula
2	1837	34.3	709	6 Q8HZ76	Q8hz76 ovis aries
3	1622	30.3	342	6 Q863B6	Q863e6 sus scrofa
4	778	14.5	1265	4 Q9P2G7	Q9p2g7 homo sapien
5	777.5	14.5	1262	4 Q8N3U0	Q8n3u0 homo sapien
6	758	14.1	2210	4 Q723E7	Q7z3e7 homo sapien
7	711.5	13.3	1417	4 Q9NYF4	Q9nyf4 homo sapien
8	711.5	13.3	1581	4 Q9BW93	Q9bw93 homo sapien
9	711.5	13.3	1787	4 Q9UP80	Q9up80 homo sapien
10	707	13.2	1338	4 Q9Y4C1	Q9y4c1 homo sapien
11	696.5	13.0	1212	4 Q9P0K0	Q9p0k0 homo sapien
12	641.5	12.0	749	11 Q8K2K4	Q8k2k4 mus musculus
13	639.5	11.9	744	11 Q8K2J6	Q8k2j6 mus musculus
14	599.5	11.2	759	4 Q9BVH6	Q9bvhe homo sapien
15	560	10.5	1192	4 Q8IY08	Q8iy08 homo sapien
16	558	10.4	854	5 Q9VHC5	Q9vhc5 drosophila

ALIGNMENTS

RESULT 1

ID	Q8WNL9	PRELIMINARY;	PRT;	1187 AA.
AC	Q8WNL9;			
DT	01-MAR-2002 (TrEMBLrel. 20, Created)			
DT	01-MAR-2002 (TrEMBLrel. 20, Last sequence update)			
DT	01-MAR-2003 (TrEMBLrel. 23, Last annotation update)			
DE	Hairless.			
OS	Macaca mulatta (Rhesus macaque).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Euthera; Primates; Catarrhini; Cercopitheciidae;			
OC	Cercopitheciinae; Macaca.			
OX	NCBI_TaxID=9544;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RA	Ahmad W., Ratterre M.S., Panteleyev A.A., Aita V.M., Sundberg J.P.,			
RA	Christiano A.M.;			
RT	"Papular Atrichia Resulting from Mutations in the Rhesus Monkey			
RT	(Macaca mulatta) hairless Gene."			
RL	Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.			
DR	EMBL; AF361864; AAL56245.1; -			
DR	InterPro; IPR003347; TF_Jmjc.			
DR	Pfam; PF02373; Jmjc_1.			
DR	SMART; SM00558; Jmjc_1.			
SQ	SEQUENCE 1187 AA; 127135 MW; D1CFED4AAA76C77E CRC64;			

Query Match 95.5%; Score 5119; DB 6; Length 1187;  
Best Local Similarity 95.3%; Pred. No. 0;  
Matches 938; Conservative 18; Mismatches 28; Indels 0; Gaps 0;

QY	1	FYIKDPSIPRLAKEPLAAEPGLFGLNSGHLORAGEAEPSLHORDGEMAGRQQNPCP	60
Db	204	FYIKDPSIPRLAKEPLAAEPGLFGLNSGHLORAGEAEPSLHORDGEMAGRQQNPCP	263
QY	61	LFLGQPDTPVWTSWPCPPGLVHTLGNVWAGPDGNLGYQLGPPATPCSPPPVPTORG	120
Db	264	LFLGQPDTPVWTSWPCPPGLVHTLGNVWAGPDGNLGYQLGPPATPCSPPPVPTORG	323
QY	121	CCSSYPPTKGGDLGPCQCEGLEGGASGESEFSEVNVKASGFRACPPSHHTLTKTWT	180

QY 481 LRHCPCQADARVWAPGAGQOKESTOKTPTPOPCSCNGDTHRTKSIKEETPDSETPAE 540  
 DB 481 LRHCPCQADARVWAPGAGQOKESTOKTPTPOPCSCNGDTHRTKSIKEETPDSETPAE 740  
 QY 541 DRAGGFLPCPSLCELLASTAVKLCIGHRIHMAFAPVTPALPSDDRTNILDSTIAQV 600  
 DB 541 DRAGGFLPCPSLCELLASTAVKLCIGHRIHMAFAPVTPALPSDDRTNILDSTIAQV 800  
 QY 601 ERKIQEALGGLRAGPGLRGLGLPLSPVRPLPPGALLWLOBPQPCRRGFHLFOEH 660  
 DB 601 ERKIQEALGGLRAGPGLRGLGLPLSPVRPLPPGALLWLOBPQPCRRGFHLFOEH 858  
 QY 661 WRQGPVLVSGIORTLQNLGTEALGALGOVQALSPLGPQPSLSGSTTFWEGFSWPE 720  
 DB 661 WRQGPVLVSGIORTLQNLGTEALGALGOVQALSPLGPQPSLSGSTTFWEGFSWPE 918  
 QY 721 LRPKSDESVLLHRLGDEDTSRVENLAASLPPEYCALHGKLNLSYLPPLGALRPLE 780  
 DB 721 LRPKSDESVLLHRLGDEDTSRVENLAASLPPEYCALHGKLNLSYLPPLGALRPLE 978  
 QY 781 POLAAAYGVSPHGHGLTKNLCVEVADIVLSILVHADTPLPAWHRAQKDFLSGLDGEGLMS 840  
 DB 781 POLAAAYGVSPHGHGLTKNLCVEVADIVLSILVHADTPLPAWHRAQKDFLSGLDGEGLMS 1038  
 QY 841 PGSOVSTVWHVFRADQAOIRREFLQMVCPAGAGALEPAGSCYLDAGLRRLREMGVS 900  
 DB 841 PGSOVSTVWHVFRADQAOIRREFLQMVCPAGAGALEPAGSCYLDAGLRRLREMGVS 1098  
 QY 901 CWTLLQAPGEAVLPAGAPHQVQGLVSTVSTVTHFSLPETSALSQALCHQGPSLPPDCHL 960  
 DB 901 CWTLLQAPGEAVLPAGAPHQVQGLVSTVSTVTHFSLPETSALSQALCHQGPSLPPDCHL 1158  
 QY 961 LYAQMWAQVFAVAVGTLQEA 984  
 DB 961 LYAQMWAQVFAVAVGTLQEA 1182  
 RESULT 2  
 S28499  
 probable finger protein - rat  
 C:Species: Rattus norvegicus (Norway rat)  
 C:Date: 12-Mar-1993 #sequence\_revision 12-Mar-1993 #text\_change 08-Dec-2000  
 C:Accession: S28499  
 R:Hoog, C.; Schalling, M.; Grunder-Brundell, E.; Daneholt, B.  
 submitted to the EMBL Data Library, June 1991  
 A:Description: Analysis of a murine germ cell-specific transcript that encodes a putative  
 A:Reference number: S28499  
 A:Molecule type: mRNA  
 A:Residues: 1-1214 <HOO>  
 A:Cross-references: EMBL:X5993; NID:q57503; PIDN:CAA42610.1; PID:q57504  
 A:Experimental source: strain Sprague Dawley  
 C:Keywords: DNA binding; zinc finger

Query Match 13.0%; Score 694.5; DB 2: Length 1214;  
 Best Local Similarity 25.8%; Pred. No. 2.8e-33;  
 Matches 260; Conservative 159; Mismatches 397; Indels 192; Gaps 40;  
 QY 113 EPPVTRGCCSSYPPTKG-GDL-----GPCKQCBEGLEGASGASPESEVNKASGRACRP 167  
 DB 244 KPDVCKAGLLSSKSSQVAGDLKILSEPKGSC-----IQPKTNTDQESRLESTP 292  
 QY 168 PSHTTKLKTWLTTHSEGFCEPFG-CPEVERBPVRLRALKRAS-----PEVQGA- 217  
 DB 293 QPVTGLTKCELVTKTSSKAEDNATAPELQKR-----LEHTASTPDGLSDKPEVEAGV 345  
 QY 218 --MGSPAPKRPDPFGTAEGAGQGWQVRVTSIGNKDVDS---GQHDQKGPQDQASL 272  
 DB 346 TRINSCSEKK-----VGPSDLGSSQ-QMLKTSV-KVDHDSCTRTSSNKTTPPARKSVL 398  
 QY 273 QDPG-----LQDIFCLALPAKLAQCSQCA-----QAAGEGGGHACHSQOVR 313  
 DB 399 TDPDKLKLQSGEAFVQDDSCVNIQALPKRCRELDLSLKDQKDSFVFCFFPFR 458

QY 314 RSPILG--GELQOEBEDTATNSSSEBPG-----SGFDSRLSTGLAKHLLSLGLDRLC 362  
 DB 459 RLQFNKHLVLRVEGFLTPNKYDSEALGLWLPLTKNVVGTD--LDT--AKYILANIGDHFC 514  
 QY 363 RLLRRERAL-----AWAQREGQPAVGTGSPGIPRCCSRCHHGLFNTHWRCPKCS 413  
 DB 515 QMWIASEKEAMSTIEPHRQVAMKR-----AV-----KGVREMCVCDTTFIFNLHVMVCPRCG 564  
 QY 414 HRLVCACGRVAGTGRAREKAGFQOSABECTQEAHGAACSLMLTQFVSSQALAEILSTAMH 473  
 DB 565 FGVCVDCVRLX-RKNCQCAAYKTFSWIRCVKSQIHEPENLMPQIIPCKALYDVGDIVH 623  
 QY 474 QVWVKFDIRHCPQADARVWAPGAGQOKESTOKTPTPOPCSCNGDTHRTKSIKEETPD 533  
 DB 624 SVRAKGIKANCPC-SNRQFKLFSKPAKEDLKQASLSGEKPSLGTVMVQSSVLEPAVA 682  
 QY 534 SAETPAEDRAGRGLPCPS-----LCCELLASTAVK-----LCLGHRIHMAFAPVTP 580  
 DB 683 CGEAPSPASNVKPI-CPANTSPNLWLADLTSGNVNKENKEKQLTMPILKNEIKCLPPLP 741  
 QY 581 ALPSDDRTITNILDSTIAQVVE-----RKIQEALGP-----GLRAGPGL----- 619  
 DB 742 PLNKSSTVILHTFNSTILTPTVSNNSGFLNLLNSTGKTENGLKNTKILDDIFASLVQN 801  
 QY 620 -----RKGLGLPLSPVRPLPPP-----GALLWLOBPQPCRRGFHLFOEHWRQSQP 666  
 DB 802 KTSDDLKSRPQGLTIKPSILGFTDPHYWLCDNRLCLQDPN--NKSNNVNFRECWKQSQP 859  
 QY 667 VLVSGIORTLQNLGTEALGALGOVQALSPLGPQPSLSGSTT--FWEGFSWPELRPK 724  
 DB 860 VMVSGVHHKLNTELWKPESFRKEFGEQEVLDVNCRTNEIITGATVGVDFWDGFDPNRLK 919  
 QY 725 SD---EGSVLLHRLGDE-----TSRVENLAASLPPEYCALHGKLNLSYLPPLGAL 776  
 DB 920 NEKEKEPWLKLDWPPGEDFRDMPSRFDLMANIPLEPYTRDRDGLKLNLSRLEPNFVR 979  
 QY 777 RPLFQPLWAAYG-VSPHRHGLTKNLCVEVADIVLSILVHADTPLPAWHRAQKDFLSGLD 835  
 DB 980 PDLGPKMAYAGLITPEDRKYGTTNLHLDVSDAANVMVVGIPKGCQCEBEEVLRITQD 1039  
 QY 836 -----EGLWSGSSQSVTVWHVFRADQAOIRREFLQMV-----PAGAGALEP 877  
 DB 1040 DSDELTKRFTIEGKEKPGA-----LWHYAAKDTKIREFLKCKVSEEQGENPADH---DP 1092  
 QY 878 GAPGSCYLDAGLRRLREMGVSCWTLILQAPGEAVLPAGAPHQVQGLVSTVSTVTHFSL 937  
 DB 1093 IHDQSWILDRSLRKLKQYGVQGWAIQVPLGVDVVFIPAGAPHQVHNLVYSCIKVAEDFVS 1152  
 QY 938 PETA-----LSAQLCHQGPSLPPDCHLLYQMDWAVFOAVKAVGTLQ 981  
 DB 1153 PEHVKHCFLWTQEPRLH--SQTHTNHEDKLQVKNVIYHVAVKDAVAMLK 1198

## RESULT 3

D86254  
 hypothetical protein [imported] - Arabidopsis thaliana  
 C:Species: Arabidopsis thaliana (mouse-ear cress)  
 C:Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 31-Mar-2001  
 C:Accession: D86254  
 R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,  
 ansen, N.F.; Hughes, B.; Huizart, L.  
 Nature 408, 816-820, 2000  
 A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.;  
 C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziani,  
 Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.  
 A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon, I.  
 ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.  
 A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.  
 A:Reference number: A86141; MUID:21016719; PMID:11130712  
 A:Accession: D86254  
 A>Status: preliminary

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model  
Run on: June 29, 2004, 10:24:04 ; Search time 27 Seconds  
(without alignments)  
3505.647 Million cell updates/sec

Title: US-10-024-368-2  
Perfect score: 5358  
Sequence: 1 FFKDPSIPRLAKEPLAAAE.....MDWAVQAVKVAAGTLQBAK 984

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR 78:\*  
1: pir1:\*  
2: pir2:\*  
3: pir3:\*  
4: pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	4102.5	76.6	1182	2 I48378	hairless protein - mouse
2	694.5	13.0	1214	2 S28499	probable finger pr
3	335	6.3	851	2 D86254	hypothetical prote
4	312.5	5.8	906	2 T01440	hypothetical prote
5	247	4.6	1701	2 T43213	ENBP1 protein - ba
6	239.5	4.5	1629	2 T06461	DNA-binding protei
7	239	4.5	1641	2 T10955	early nodulin bind
8	233.5	4.4	730	2 B85013	hypothetical prote
9	233.5	4.4	730	2 T10539	hypothetical prote
10	205	3.8	1616	2 T13183	gene APX1 protein
11	199	3.7	1106	2 JQ0405	hypothetical 119.5
12	197.5	3.7	2282	2 T42717	DNA-binding protei
13	190.5	3.6	660	1 Q0BE3	BHLFI protein - hu
14	190	3.5	728	2 T05151	hypothetical prote
15	184.5	3.4	1207	2 T00378	KIAA0641 protein -
16	183	3.4	1436	1 A45344	immediate-early pr
17	182.5	3.4	1433	2 A46053	bulbosus pemphigoid
18	182.5	3.4	3942	2 T42730	Bassoon protein -
19	180	3.4	1733	1 B45344	probable nuclear a
20	179.5	3.4	924	2 S27923	gene LF3 protein -
21	178	3.3	677	2 S23296	collagen alpha 2(I
22	177.5	3.3	1736	2 T00391	hypothetical prote
23	177	3.3	711	2 S05381	VGFA protein prec
24	176	3.3	2774	2 A43359	microtubule-associ
25	174.5	3.3	3938	2 T42761	Bassoon protein -
26	173.5	3.2	1146	2 A38587	collagen, cornea-s
27	172	3.2	950	2 F86222	hypothetical prote
28	171.5	3.2	2142	2 B35098	MHC class III hist
29	170.5	3.2	4957	2 T03455	ALR protein - huma

30 170.5 3.2 5262 2 T03454 ALR protein - huma  
31 170 3.2 1670 1 CGHU3B collagen alpha 3(I  
32 168.5 3.1 1466 1 CGHU7L collagen alpha 1(I  
33 168 3.1 1317 2 T03748 apoptosis associat  
34 167.5 3.1 1300 2 T03166 probable immediate  
35 166.5 3.1 1142 2 JX0369 collagen alpha 1(X  
36 165.5 3.1 1419 2 A41182 collagen alpha 1(I  
37 165.5 3.1 1460 1 EDBE1F immediate-early pr  
38 165.5 3.1 1487 2 B41182 collagen alpha 1(I  
39 164 3.1 1464 2 S59856 collagen alpha 1(I  
40 163.5 3.1 940 2 JE0291 FB19 protein - hum  
41 163.5 3.1 1373 1 A43291 collagen alpha 2(I  
42 163.5 3.1 1418 2 T45467 collagen alpha 1(I  
43 163 3.0 1958 2 B40505 hypothetical prote  
44 162 3.0 1366 1 CGHU2S collagen alpha 2(I  
45 162 3.0 1414 1 S23809 collagen alpha 2(I

ALIGNMENTS

RESULT 1

I48378  
hairless protein - mouse  
C:Species: Mus musculus (house mouse)  
C:Date: 02-Jul-1996 #sequence\_revision 02-Jul-1996 #text\_change 05-Nov-1999  
C:Accession: I48378  
R:Cachon-Gonzalez, M.B.; Fenner, S.; Coffin, J.M.; Moran, C.; Best, S.; Stoye, J.P.  
Proc. Natl. Acad. Sci. U.S.A. 91, 7717-7721, 1994  
#Title: Structure and expression of the hairless gene of mice.  
#Reference number: I48378; MUID:94329587; PMID:8052649  
#Accession: I48378  
#Status: preliminary;  
#Molecule type: mRNA  
#Residues: 1-1182 <RES>  
#Cross-references: EMBL:Z32675; NID:9531706; PIDN:CAA83587.1; PID:9531707

Query Match 76.6%; Score 4102.5; DB 2; Length 1182;  
Best Local Similarity 78.4%; Pred. No. 2.2e-234;  
Matches 771; Conservative 62; Mismatches 144; Indels 7; Gaps 5;

QY	1	FFKDP	SIPRLAKEPLAAAE	PGLFGLNSG	HLORACEA	ERPSLH	RDGEMG	AGRQ	NCPCP	60
DB	206	FHKD	PNILRLAKEPL	--AESG	MLGLAPG	GHLLQQA	CESEGPS	LHQD	GETGAGRQ	263
QY	61	LFLG	QDTPVPT	SWPAC	PPGLVHT	LGNWAG	PGDGNL	GYQLG	PPATPR	120
DB	264	VELG	YDTPVPT	APWPC	PCFGLVH	SLGNWAG	PGSNLS	GYQLG	PPATPR	323
QY	121	CCSS	YPTTKG	DLPCK	QCGK	QEGLEG	ASGASE	SEEVNK	ASGPAC	180
DB	324	CCSS	HLPAR	EGDLG	PCRKQ	QDSPEG	SGSGP	GESSEERN	KADS-RAC	382
QY	181	RISE	QECPRG	CEVEE	RPVAR	LRAK	RAGSP	EVQAMG	SPAPK	240
DB	383	RHSE	QECPCG	CGSGKE	ESATGL	RALK	RAGSP	EVQASR	GPAPR	442
QY	241	WQSV	RTSIGN	KDVDS	GQHQD	EQKQD	GQASL	QDIP	CLALPA	300
DB	443	WQTP	PTII	IGSK-AEA	EQEQE	QGRPRD	GRIRL	QESL	VDTS	501
QY	301	EGG	HACH	QQVRR	SPLGGE	LQOE	EDTAN	NSSE	EGPG	360
DB	502	EVGL	TGH	SQKSR	SPL-EK	QLEED	SSAT	SEGG	GGPG	560
QY	361	LCFL	IRRE	REAL	AWARE	QCGP	AVTGD	SPGIP	CCSR	420
DB	561	LCFL	IRRE	REAL	AWARE	QCGP	AVTGD	SPGIP	CCSR	620
QY	421	GRV	AGTGR	AREK	AGFQ	EQSAE	CTQEA	GHAA	CSMLT	480
DB	621	GRI	AGAK	NRK	ETG	SQEQ	ETDD	CAQ	EAGH	680



```

QY 121 CCSYPTTKGGDLGPGCGKQEGLEGASGASPSSEVNKASGPRACPPSHHTKLKKTWLT 180
Db 352 CCSSHLPAEGDGPFCRKQDSPEGSSGPGSSSEERNKA-GSRASPTSHHTKLKKTWLT 410
QY 181 RHSEQFECPRGCPVEVERPVARLRALKRAGSPEVQGMGSPAPKRPDPFPPTAQGAGG 240
Db 411 RHSEQFECGCGPGKGESPATGLRALKRAGSPEVQAGG-PAPKRPSTHTFPPTGROGARA 469
QY 241 WQEVRTDTSIGNKDVDSGHDGKQPDGQASLQDPCGLQDIPCLALPAKLAQCSCCAQAAG 300
Db 470 WQETPTSTGSK-AEAQQQEEQGRGDRIRLRRESLEDTSQHLHLAGVTQCPCSCVQAAG 528
QY 301 EGGGHACHSQVRRSPGLGQEQEEDTATNSSSEEGPGSPDRLSTGLAKHLLSLGDR 360
Db 529 EVEILTSHSQSHKLPLEBKPLEEDSCAT--SEEGGSSPEASINKGLAKHLLSLGDR 585
QY 361 LCRLLRERREALAWAQREGQGAFAVGDSPGIPRCSRCHHGLFNTHWRCPRCSHRLCVAC 420
Db 586 LCRLLRERREALAWAQREGQGAFAVGDSPGIPRCSRCHHGLFNTHWRCPRCSHRLCVAC 645
QY 421 GRVAGTGRAREKAGFOEQSAECTOBAGHAACSLMLTQFVSSQALAEELSTAMHQQVWKFD 480
Db 646 GRIAGAKNREKTSREQRTDCAQEAAGHAACSLMLTQFVSSQALAEELSTAMHQQVWKFD 705
QY 481 IRGHCPQADARWAPGDAGQOKESTQKTPPTPQPSGNGDTHRTKSIKEETPDSPAETPAE 540
Db 706 IRGHCFQCVDAWAPGDGQOKKEPTKTPPAPQLSCNGDSNRKDKIKEETPDSTESPAE 765
QY 541 DRAGRGPLPCPSLCCELLASTAVKLCGLHERIHMAFAPVTPALPDSDDRIITNILDSTIAQV 600
Db 766 DRAGRSPLPCPSLCCELLASTAVKLCGLHERIHMAFAPVTPALPDSDDRIITNILDSTIAQV 825
QY 601 ERKIQEKALGPGLRAGPGRKGLGLPLSPVRPLPPGALLMLQEPQPCPRGFGHLPQSH 660
Db 826 ERKIQEKALGPGLRAGPGRKGLGLPLSPVRPLPPGALLMLQEP--PKHGFRFLQEH 883
QY 661 WRQGPVLVSGIQTILQGNLWTEALGALGGVQVQALSPGPPQPSLGSSTTWEGFSWPE 720
Db 884 WRQGPVLVSGIQTILRLSLMGMEALTLGGQVQVTLTALGPQPTSLDSTAFWKGFSPHE 943
QY 721 LRPKDEGVLILHRAIGDETSRVENLAASLPLPYCALHGKINLASVLPGLALRPLE 780
Db 944 ASPKLDEGVLILHRAIGDETSRVENLAASLPLPYCALHGKINLASVLPGLALRPLE 1003
QY 781 POLMAAYGVSPHRHGLGTNKLVEVADLVSLVHADTLPWHAQKDFLSGLDGEGLWS 840
Db 1004 POLMAAYGVNSRHGLGTNKLVEVSDLSILVHAEAQLPWVRAQKDFLSGLDGEGLWS 1063
QY 841 PGSQVSTVWHVFRQAQRIIRFLQVCPAGAGALEPGAGSCYLDAGLRRLRRLREWGV 900
Db 1064 PGSQVSTVWHVFRQAQRIIRFLQVCPAGAGALEPGAGSCYLDAGLRRLRRLREWGV 1123
QY 901 CWTLLQAPGEAVLPAGAPHQVGLVSTVTOHFLSPETSALSQQLCHQGPSLPDCHL 960
Db 1124 CWTLLQAPGEAVLPAGAPHQVGLVSTVTOHFLSPETSALSQQLCHQGPSLPDCHL 1183
QY 961 LYAQMDDWAFQAVKAVGTLQEAQ 984
Db 1184 LYAQMDDWAFQAVKAVGTLQEAQ 1207

RESULT 14
ABU62543
ID ABU62543 standard; protein; 1207 AA.
XX AC ABU62543;
XX DT 08-SEP-2003 (first entry)
XX DE Rat Hairless (Hr) polypeptide.
XX KW Rat; hairless; Hr; thyroid hormone receptor; hair growth; hair loss;
KW family pedigree; radiation hybrid; somatic cell hybrid.

```

```

XX Rattus sp.
XX US2003027300-A1.
XX 06-FEB-2003.
XX 21-DEC-2001; 2001US-00024368.
XX 07-APR-1998; 98US-0080888P.
XX 07-APR-1999; 99US-00287354.
XX (CARN-) CARNEGIE INST WASHINGTON.
XX Thompson CC;
XX WPI; 2003-492034/58.
XX New Hairless polypeptide and polynucleotide, useful for identifying and
XX detecting this genetic marker in family pedigrees or human-rodent somatic
XX cell hybrids, or detecting interacting proteins that bind hairless gene
XX or protein.
XX Example; Fig 1; 34pp; English.
XX The invention relates to a human hairless (HR) polynucleotide encoding
XX the hairless (Hr) polypeptide. The invention also relates to an
XX expression system comprising an expression construct which produces a
XX polypeptide with hairless transcription factor activity, a reporter
XX construct comprising a transcription regulatory region responsive to
XX hairless transcription activity to regulate transcription of the reporter
XX gene which is mediated by the transcription regulatory region and methods
XX of screening for chemical agents which modulate hairless-mediated
XX transcription, binding between hairless and thyroid hormone receptor or
XX hairless activity. The polynucleotide is useful as a probe or primer to
XX quantitate cognate RNA and DNA within cells, which can be subsequently
XX used to correlate hair growth or loss with hairless expression or
XX hairless-regulated transcription. The amino acid sequence of hairless
XX antigen can be used for preparing specific binding molecules (e.g.
XX polyclonal or monoclonal antibodies) for monitoring protein expression,
XX for affinity purification and for functional studies. The human hairless
XX polynucleotide, polypeptide or specific binding molecule may be used to
XX identify and detect this genetic marker in family pedigrees, radiation
XX hybrids or human-rodent somatic cell hybrids, and as an affinity tag to
XX identify, isolate and detect interacting proteins that bind the hairless
XX gene or protein. This sequence represents a rat hairless polypeptide of
XX the invention
XX SQ Sequence 1207 AA;
XX Query Match 75.9%; Score 4067; DB 6; Length 1207;
XX Best Local Similarity 77.7%; Pred. No. 1.8e-298;
XX Matches 765; Conservative 65; Mismatches 146; Indels 8; Gaps 5;
QY 1 FYYKDSIPLAKEPIAAAEPLGLNSGHLORAGAEERPSLHQRDGEMGAGROQNP 60
Db 232 FYHKDPNILLPAKEPIAAASEGMLAPGQHLOQACDAEGPSLHQRDGETGACRQQLCP 291
QY 61 LFLGQPDTPVWTSWPACPPGLVHTLGNVWAGDGNLGYQLGPPATPCRPCSPPEVTVQ 120
Db 292 VFLGYDTPVTPWSPCPPELGNVWAGPGNSFGYQLGPPVTPRCSPGPPTPPG 351
QY 121 CCSSYPTTKGGDLGPGCGKQEGLEGASGASPSSEVNKASGPRACPPSHHTKLKKTWLT 180
Db 352 CCSSHLPAEGDGPFCRKQDSPEGSSGPGSSSEERNKA-GSRASPTSHHTKLKKTWLT 410
QY 181 RHSEQFECPRGCPVEVERPVARLRALKRAGSPEVQGMGSPAPKRPDPFPPTAQGAGG 240
Db 411 RHSEQFECGCGPGKGESPATGLRALKRAGSPEVQAGG-PAPKRPSTHTFPPTGROGARA 469
QY 241 WQEVRTDTSIGNKDVDSGHDGKQPDGQASLQDPCGLQDIPCLALPAKLAQCSCCAQAAG 300
Db 470 WQETPTSTGSK-AEAQQQEEQGRGDRIRLRRESLEDTSQHLHLAGVTQCPCSCVQAAG 528

```



CC quantitate cognate RNA and DNA within cells, which can be subsequently  
 CC used to correlate hair growth or loss with hairless expression or  
 CC hairless-regulated transcription. The amino acid sequence of hairless  
 CC antigen can be used for preparing specific binding molecules (e.g.  
 CC polyclonal or monoclonal antibodies) for monitoring protein expression,  
 CC for affinity purification and for functional studies. The human hairless  
 CC polynucleotide, polypeptide or specific binding molecule may be used to  
 CC identify and detect this genetic marker in family pedigrees, radiation  
 CC hybrids or human-rodent somatic cell hybrids, and as an affinity tag to  
 CC identify, isolate and detect interacting proteins that bind the hairless  
 CC gene or protein. This sequence represents a mouse hairless polypeptide of  
 CC the invention  
 XX  
 SQ Sequence 1182 AA;

Query Match 76.6%; Score 4102.5; DB 6; Length 1182;  
 Best Local Similarity 78.4%; Pred. No. 3.5e-301;  
 Matches 771; Conservative 62; Mismatches 144; Indels 7; Gaps 5;

QY 1 FYKDPSTPLRAKEPLAAEPCLEGLNSGHLORAGEAERPSLHORDEMGAGROQPCP 60  
 DB 206 FYKDPSTPLRAKEPLAAEPCLEGLNSGHLORAGEAERPSLHORDEMGAGROQPCP 263

QY 61 LFLGQDTPVPTWSPACPGVLVHTLGNVWAGPGDGNLGYQLGPPATPCPPEPPTQRG 120  
 DB 264 VFLGQDTPVPTWSPACPGVLVHTLGNVWAGPGDGNLGYQLGPPATPCPPEPPTQRG 323

QY 121 CCSSYPPTKGGDLGPGCGKQCGLEGAGAGSEPEVNKASGPRACPSHHTKLKTLT 180  
 DB 324 CCSSHLPAREGLDPCRCQDPSGEGSGGSESEERNKADS-RACPPSHHTKLKTLT 382

QY 181 RSHQFECPCPCPEVEERPARLRALKAGSPEVQMGSPAPKPPDPFGTAEOGAGG 240  
 DB 383 RSHQFECPCPCPEVEERPARLRALKAGSPEVQMGSPAPKPPDPFGTAEOGAGG 442

QY 241 WQEVRTDSIGNKVDGSDHDKQPGDQASLQDIPCLALPAKLAOCQCAQAG 300  
 DB 443 WQETPTIIGSK-AEAEOQEGRQPRDRILOESRLVDTSCQHLAGVTQCCVQAG 501

QY 301 EGGHACHSQVRSPLGEGLOQEDTATNSSSEGGPGSDRLSTGLAKHLISGLGDR 360  
 DB 502 EVGLVTHGSKSRSPLE-EEKQLEEDSATSSEGGGPGPEASLNKLAHLLSGLGDR 560

QY 361 LCRLLRERREALWAQREGQPAVTPGPGIPCCSRCHGLFNTHWRCPSHRLCVAC 420  
 DB 561 LCRLLRERREALWAQREGQPAVTPGPGIPCCSRCHGLFNTHWRCPSHRLCVAC 620

QY 421 GRVAGTGRAREKAGFQEQSABECTQEAHAACSLMLTQFVSSQALAEISLTHAMHVKFD 480  
 DB 621 GRIAGAKNREKTSQEQHTDDCAQEAHAACSLMLTQFVSSQALAEISLTHAMHVKFD 680

QY 481 IRGHCPCQADARVWAPGAGQOKSTQKTPPTQPSGNDTHRTKSIKEETDPSAETPAE 540  
 DB 681 IRGHCPCQADARVWAPGAGQOKSTQKTPPTQPSGNDTHRTKSIKEETDPSAETPAE 740

QY 541 DRAGGRLPCSLCELLASTAVKLCIHERIHMFAFVTPALPSDDRTNILDSTIAQV 600  
 DB 741 DRAGGRLPCSLCELLASTAVKLCIHERIHMFAFVTPALPSDDRTNILDSTIAQV 800

QY 601 ERKIOEKALGFLRAGPGLRGLGLPLSFVRPLPFPFGALLWLPQPCRRGPHLFOEH 660  
 DB 801 ERKIOEKALGFLRAGPGLRGLGLPLSFVRPLPFPFGALLWLPQPCRRGPHLFOEH 858

QY 661 WRQGPVLVSGIORTLQNLWGTEALGALGGVQVQLSPGPPQPSLSTGTFWGFSPWE 720  
 DB 859 WRQGPVLVSGIORTLQNLWGTEALGALGGVQVQLSPGPPQPSLSTGTFWGFSPWE 918

QY 721 LRPKSDGSLVLLHREALGDEDTSRVENLAASIPLEYCALHCKNLASVLPGLALRPLE 780  
 DB 919 LRPKSDGSLVLLHREALGDEDTSRVENLAASIPLEYCALHCKNLASVLPGLALRPLE 978

QY 781 POLMAAGVSPHRLGTPKNCLEVAIDLVSILVHADTPLPAWHRAQDKFLSGLDGEGLWS 840

DB 979 POLMAAGVSPHRLGTPKNCLEVAIDLVSILVHAEALPWPYRAQDKFLSGLDGEGLWS 1038  
 QY 841 PGSOVSTVWHVFRQAQDAQRIIRRFQVWCPAGAGALEPCGAPGSCYLDAGLRRRLREMGVS 900  
 DB 1039 PGSOVSTVWHVFRQAQDAQRIIRRFQVWCPAGAGALEPCGAPGSCYLDAGLRRRLREMGVS 1098  
 QY 901 CWTLLQAPGEAVLVPAGAPHOVQGLVSTVSTVTOHFLSPETSAQLCHQGPSLPPDCHL 960  
 DB 1099 CWTLLQAPGEAVLVPAGAPHOVQGLVSTVSTVTOHFLSPETSAQLCHQGPSLPPDCHL 1158  
 QY 961 LYAQMWAQVFAOVAVGVTLOEAK 984  
 DB 1159 LYAQMWAQVFAOVAVGVTLOEAK 1182

RESULT 13  
 AAEL19797  
 ID AAEL19797 standard; protein; 1207 AA.  
 XX  
 AC AAEL19797;  
 XX  
 DT 18-JUN-2002 (first entry)  
 XX  
 DE Rat Hairless protein (Hr).  
 XX  
 KW Rat; Hairless protein; Hr protein; HR gene; dermatological condition;  
 KW hair loss; gene therapy.  
 XX  
 OS Rattus norvegicus.  
 XX  
 FN US6348348-B1.  
 XX  
 PD 19-FEB-2002.  
 XX  
 PF 07-APR-1999; 99US-00287354.  
 XX  
 PR 07-APR-1998; 98US-0080888P.  
 XX  
 PA (CARN-) CARNEGIE INST WASHINGTON.  
 XX  
 PI Thompson CC;  
 XX  
 DR WPI; 2002-204622/26.  
 XX  
 PT Novel expression construct, useful in the diagnosis and treatment of  
 PT dermatological conditions, such as hair loss, contains a Hairless gene  
 PT sequence.  
 XX  
 PS Example; Fig 1; 48pp; English.  
 XX  
 CC The invention relates to human Hairless (Hr) polypeptides and nucleic  
 CC acid molecules (HR) encoding such polypeptides. The invention also  
 CC relates to the Hairless expression constructs which may be used in  
 CC transcription assays. Sequences of the invention are used in diagnosis  
 CC and treatment of dermatological conditions such as hair loss. They are  
 CC also used in gene therapy. Polynucleotides of the invention can be used  
 CC as probes for the detection of hair loss. The present sequence is rat  
 CC Hairless protein (Hr)  
 XX  
 SQ Sequence 1207 AA;

Query Match 75.9%; Score 4067; DB 5; Length 1207;  
 Best Local Similarity 77.7%; Pred. No. 1.8e-298;  
 Matches 765; Conservative 65; Mismatches 146; Indels 8; Gaps 5;

QY 1 FYKDPSTPLRAKEPLAAEPCLEGLNSGHLORAGEAERPSLHORDEMGAGROQPCP 60  
 DB 232 FYKDPSTPLRAKEPLAAEPCLEGLNSGHLORAGEAERPSLHORDEMGAGROQPCP 291

QY 61 LFLGQDTPVPTWSPACPGVLVHTLGNVWAGPGDGNLGYQLGPPATPCPPEPPTQRG 120  
 DB 292 VFLGQDTPVPTWSPACPGVLVHTLGNVWAGPGDGNLGYQLGPPATPCPPEPPTQRG 351

XX	Query Match	76.6%;	Score 4102.5;	DB 5;	Length 1182;	
PR	Best Local Similarity	78.4%;	Pred. No. 3.5e-301;			
XX	Matches 771;	Conservative 62;	Mismatches 144;	Indels 7;	Gaps 5;	
PA	1 FYYKPSIPRLAKEPIAAEPGLFGLNSGHLQAGEAERPSLHQRDGEAGRQONPCP	60				
XX	206 FYHKDPNILLPAKEPL--AESGMLGLAPGCHLQACESEGPSLHQRDGETGAGRQONLCP	263				
XX	61 LFLGQDDTVPTWSPACPLGVLTGLNHWAGPGDGMGLYGLGPPATPCPSPPVPTQRG	120				
XX	264 VFLGYDTPVPRAPWPCPLGVLSLGNHWAGPGSNLSYGLGPPATPCPSPPVPTQRG	323				
XX	121 CCSSYPTKGGDLGPGCKQEGLEGASGASRSEVKNASGPRACPPSHHTKLKKTWLT	180				
XX	324 CCSSHLPAREGDLGCRKQDSPEGGSSGPGSESEERNKADS--RACPPSHHTKLKKTWLT	382				
XX	181 RHSEQFECPRCPEVEERPVARLRALKRAGSPVQVAMGSPAPKRPDPFPFGTAEGAGG	240				
XX	383 RHSEQFECPGCGCKEESATGLRALKRAGSPVQVAMGSPAPKRPDPFPFGTGRQARA	442				
XX	241 WQVVRTSTGNKVDGSHQDEKGPQDQASLQDGLDIPCLALPAKLAQCSQAQAG	300				
XX	443 WQETPTIIGSK-AAEABQEQEQRGPRDGRIRLQESRLVDTSCQHILAGVTQCQSCVQAAG	501				
XX	301 EGGGHACHSQVRRSPILGGELQOEEDTATNSSEEGPGSPDSRLSTGLAKHLLSLGLDR	360				
XX	502 EVGLTGHQSKRRSPL--EKQLEEDSSATSEGGGPGPEASLNKGLAKHLLSLGLDR	560				
XX	361 LCRLLRREALAWAQREGQPAVTDGSPGIPRCCSRCHGLFNTHWRCPRCSHRLCVAC	420				
XX	561 LCRLLRREALAWAQREGQPAVTDGSPGIPRCCSRCHGLFNTHWRCPRCSHRLCVAC	620				
XX	421 GRVAGTGRAREKAGFQESNEECTQAGHACSLMTQFVSSQALAEALSTAMHQQVWKF	480				
XX	621 GRITAGAGRKREKTSQEQFHTDDCAEAGHAGAACSLILTFVSSQALAEALSTAMHQQVWKF	680				
XX	481 IRGHCFQADARVWAPGDAGQCKESQKTPPTQPSGNGDTHRTKSIKEETPDQSAETPAE	540				
XX	681 IRGHCFQADARVWAPGDAGQCKESQKTPPTQPSGNGDTHRTKSIKEETPDQSAETPAE	740				
XX	541 DRAGRGPLPCPSICELLASTAVKLCIGHERIHMAFAPVTPALPSDDRTNILDIIIAQVV	600				
XX	741 DGAGRGPLPCPSICELLASTAVKLCIGHERIHMAFAPVTPALPSDDRTNILDIIIAQVV	800				
XX	601 ERKIQKALPGIRAGPLRKGLPLSPVPRPLPPGALLMLQEQPCPRRPFHFOEH	660				
XX	801 ERKIQKALPGIRAGPLRKGLPLSPVPRPLPPGALLMLQEQPCPRRPFHFOEH	858				

QY	661	WRQGPVLVSGIQRITLQGNLWGTALGQVQALSPGPPQPSLSGTTFWEGFSWPE	720
DB	859	WRQGPVLVSGIQRITLQGNLWGTALGQVQALSPGPPQPSLSGTTFWEGFSWPE	918
QY	721	LRPKDEGSGVLLHRLALGDEDTSRVENLAASLPLPEYCALHGKLNLSYLPGLALRPLE	780
DB	919	TRPKLDEGSGVLLHRLALGDEDTSRVENLAASLPLPEYCALHGKLNLSYLPGLALRPLE	978
QY	781	POLMAAYGVSPHGHGTLTKNLCVEVADILVSLVHADTLPAPWHAQKDFLSGLDGEGLWS	840
DB	979	POLMAAYGVSPHGHGTLTKNLCVEVADILVSLVHADTLPAPWHAQKDFLSGLDGEGLWS	1038
QY	841	PGSQSVTWVHVFRAQDAQRIRRFQWVCPAGAGALFPGAPGSCYLDAGLRRLREEMGVS	900
DB	1039	PGSQSVTWVHVFRAQDAQRIRRFQWVCPAGAGALFPGAPGSCYLDAGLRRLREEMGVS	1098
QY	901	CWTLLOAPGEAVLVPAGAPHQVQGLVSTVSTQHFSLPETSALSQAQICHQGPSLPDCHL	960
DB	1099	CWTLLOAPGEAVLVPAGAPHQVQGLVSTVSTQHFSLPETSALSQAQICHQGPSLPDCHL	1158
QY	961	LYAQMDMAVFOAVKVAVGTLOEAK	984
DB	1159	LYAQMDMAVFOAVKVAVGTLOEAK	1182
RESULT 12			
ABU62544			
ID	ABU62544	standard; protein; 1182 AA.	
XX	AC	ABU62544;	
XX	AC	AC	
DT	08-SEP-2003	(first entry)	
XX	DE	Mouse Hairless (Hr) polypeptide.	
XX	KW	Mouse; hairless; Hr; thyroid hormone receptor; hair growth; hair loss;	
XX	KW	family pedigree; radiation hybrid; somatic cell hybrid.	
OS	Mus sp.		
XX	PN	US2003027300-A1.	
XX	PD	06-FEB-2003.	
XX	PF	21-DEC-2001; 2001US-00024368.	
XX	PR	07-APR-1998; 98US-0080888P.	
XX	PR	07-APR-1999; 99US-00287354.	
XX	PA	(CARN-) CARNEGIE INST WASHINGTON.	
XX	PI	Thompson CC;	
XX	DR	WPI; 2003-492034/58.	
XX	PT	New Hairless polypeptide and polynucleotide, useful for identifying and	
XX	PT	detecting this genetic marker in family pedigrees or human-rodent somatic	
XX	PT	cell hybrids, or detecting interacting proteins that bind hairless gene	
XX	PS	or protein.	
XX	PS	Example; Fig 1; 34pp; English.	
XX	CC	The invention relates to a human hairless (HR) polynucleotide encoding	
XX	CC	the hairless (Hr) polypeptide. The invention also relates to an	
XX	CC	expression system comprising an expression construct which produces a	
XX	CC	polypeptide with hairless transcription factor activity, a reporter	
XX	CC	construct comprising a transcription regulatory region responsive to	
XX	CC	hairless transcription activity to regulate transcription of the reporter	
XX	CC	gene which is mediated by the transcription regulatory region and methods	
XX	CC	of screening for chemical agents which modulate hairless-mediated	
XX	CC	transcription, binding between hairless and thyroid hormone receptor or	
XX	CC	hairless activity. The polynucleotide is useful as a probe or primer to	

XX Human; hairless; Hr; thyroid hormone receptor; hair growth; hair loss;  
 KW family pedigree; radiation hybrid; somatic cell hybrid.  
 XX Homo sapiens.  
 XX US2003027300-A1.  
 XX 06-FEB-2003.  
 XX 21-DEC-2001; 2001US-00024368.  
 XX 07-APR-1998; 98US-0080888P.  
 XX 07-APR-1999; 99US-00287354.  
 XX (CARN-) CARNEGIE INST WASHINGTON.  
 XX Thompson CC;  
 XX WPI; 2003-492034/58.  
 XX New Hairless polypeptide and polynucleotide, useful for identifying and  
 PT detecting this genetic marker in family pedigrees or human-rodent somatic  
 PT cell hybrids, or detecting interacting proteins that bind hairless gene  
 PT or protein.  
 XX Example; Fig 1; 34pp; English.  
 XX The invention relates to a human hairless (HR) polynucleotide encoding  
 CC the hairless (Hr) polypeptide. The invention also relates to an  
 CC expression system comprising an expression construct which produces a  
 CC polypeptide with hairless transcription factor activity, a reporter  
 CC construct comprising a transcription regulatory region responsive to  
 CC hairless transcription activity to regulate transcription of the reporter  
 CC gene which is mediated by the transcription regulatory region and methods  
 CC of screening for chemical agents which modulate hairless-mediated  
 CC transcription, binding between hairless and thyroid hormone receptor or  
 CC hairless activity. The polynucleotide is useful as a probe or primer to  
 CC quantitate cognate RNA and DNA within cells, which can be subsequently  
 CC used to correlate hair growth or loss with hairless expression or  
 CC hairless-regulated transcription. The amino acid sequence of hairless  
 CC antigen can be used for preparing specific binding molecules (e.g.  
 CC polyclonal or monoclonal antibodies) for monitoring protein expression,  
 CC for affinity purification and for functional studies. The human hairless  
 CC polynucleotide, polypeptide or specific binding molecule may be used to  
 CC identify and detect this genetic marker in family pedigrees, radiation  
 CC hybrids or human-rodent somatic cell hybrids, and as an affinity tag to  
 CC identify, isolate and detect interacting proteins that bind the hairless  
 CC gene or protein. This sequence represents a human hairless polypeptide of  
 CC the invention  
 XX SQ Sequence 1189 AA;

Query Match 98.7%; Score 5290; DB 6; Length 1189;  
 Best Local Similarity 99.0%; Pred. No. 0;  
 Matches 974; Conservative 1; Mismatches 9; Indels 0; Gaps 0;  
 QY 1 FYYKDPISIPRLAKEPLAAEEFGLFGLNSGGHQLRAGEAERPSLHORCEMGAGQQNPPCP 60  
 DB 206 FYYKDPISIPRLAKEPLAAEEFGLFGLNSGGHQLRAGEAERPSLHORCEMGAGQQNPPCP 265  
 QY 61 LFLGQPDTPVMTSWPACPPGLVHTLGNVWAGPGDGNLGYQLGPATRCPSPEPVTQRG 120  
 DB 266 LFLGQPDTPVMTSWPACPPGLVHTLGNVWAGPGDGNLGYQLGPATRCPSPEPVTQRG 325  
 QY 121 CCSSYPPTKGGDLGPGCGKQCGLEGGASGSEPEEVNKGASGRACPPSHHTLKKTKWLT 180  
 DB 326 CCSSYPPTKGGDLGPGCGKQCGLEGGASGSEPEEVNKGASGRACPPSHHTLKKTKWLT 385  
 QY 181 RHSQFCPCRCPEVERPVARLRALKRAGSPFVQAGMSPAPKRPDPDPFGTAEQAGG 240  
 DB 386 RHSQFCPCRCPEVERPVARLRALKRAGSPFVQAGMSPAPKRPDPDPFGTAEQAGG 445

QY 241 WQEVRTDSIGNKVDVSGQHDQKQPDQQAASLQDFGLQDIPCLALPAKLAQCQCAQAAG 300  
 DB 446 LQEVRTDSIGNKVDVSGQHDQKQPDQQAASLQDFGLQDIPCLALPAKLAQCQCAQAAG 505  
 QY 301 EGGGHACHSQOVRRSPGLGELQQEEDTATNSSSEEGPGSPDSRLSTGLAKHLISGLGDR 360  
 DB 506 EGGGHACHSQOVRRSPGLGELQQEEDTATNSSSEEGPGSPDSRLSTGLAKHLISGLGDR 565  
 QY 361 LCRLLRERREALAWAQREGQGFVTDSPGIPRCCSCHHGLFNTHWRCPCRSRLCVAC 420  
 DB 566 LCRLLRERREALAWAQREGQGFVTDSPGIPRCCSCHHGLFNTHWRCPCRSRLCVAC 625  
 QY 421 GRVAGTGRAREKAGFQCSABECTQBAHGAACSLMLTQFVSSQALAEIETAMHQQVWVFPD 480  
 DB 626 GRVAGTGRAREKAGFQCSABECTQBAHGAACSLMLTQFVSSQALAEIETAMHQQVWVFPD 685  
 QY 481 IRGHCPCQADARVWAPGDAGQKESSTOKTPTTPOPCNGDTHRTKSIKEETPDSAETPAE 540  
 DB 586 IRGHCPCQADARVWAPGDAGQKESSTOKTPTTPOPCNGDTHRTKSIKEETPDSAETPAE 745  
 QY 541 DRAGRGPLPCPSLCELLASTAVKLCLGHERIHMAFAPVTPALPSSDDRTITNLDIIAQVV 600  
 DB 746 DRAGRGPLPCPSLCELLASTAVKLCLGHERIHMAFAPVTPALPSSDDRTITNLDIIAQVV 805  
 QY 601 ERKIQEKALGFLRAGPLRKLGLPLSPVRPRLPPFGALLWLOEPOPCPRRGPHLFQEH 660  
 DB 806 ERKIQEKALGFLRAGPLRKLGLPLSPVRPRLPPFGALLWLOEPOPCPRRGPHLFQEH 865  
 QY 661 WRQOPVLVSGIORTLOCNLWGTALGALGQVQALSPGPPQPSLGGSTTFWEGFSWPE 720  
 DB 866 WRQOPVLVSGIORTLOCNLWGTALGALGQVQALSPGPPQPSLGGSTTFWEGFSWPE 925  
 QY 721 LRPKSDGSLVLLHRAFGDEDTSRVENLAASLPUPEYCALHGKLNLAASYLPPGLALRPLE 780  
 DB 926 LRPKSDGSLVLLHRAFGDEDTSRVENLAASLPUPEYCALHGKLNLAASYLPPGLALRPLE 985  
 QY 781 POLWAAVGVSPHGHGTLKNCVEADVLSILVHADTPLPAWHRAQKDFLSGLDGEGLWS 840  
 DB 986 POLWAAVGVSPHGHGTLKNCVEADVLSILVHADTPLPAWHRAQKDFLSGLDGEGLWS 1045  
 QY 841 PGSOVSTVWVFRADQARIRRFLOMVCPCAGAGALEPCAGSCVLDAGLRRLREEVGS 900  
 DB 1046 PGSOVSTVWVFRADQARIRRFLOMVCPCAGAGALEPCAGSCVLDAGLRRLREEVGS 1105  
 QY 901 CWTLLQAPGSAVLVPAGAPHQVQGLVSTVSTVQHFSLPETSALSAQLCHQGPSLPPDCHL 960  
 DB 1106 CWTLLQAPGSAVLVPAGAPHQVQGLVSTVSTVQHFSLPETSALSAQLCHQGPSLPPDCHL 1165  
 QY 961 LYAQMNDWAVQAVKAVVGTLOEAK 984  
 DB 1166 LYAQMNDWAVQAVKAVVGTLOEAK 1189

RESULT 11  
 AAEL19798  
 ID AAEL19798 standard; protein; 1182 AA.  
 XX  
 AC AAEL19798;  
 XX  
 XX  
 DT 18-JUN-2002 (first entry)  
 XX  
 DE Mouse Hairless protein (Hr).  
 XX  
 KW Mouse; Hairless protein; Hr protein; Hr gene; dermatological condition;  
 KW hair loss; gene therapy.  
 XX  
 OS Mus musculus.  
 XX  
 FN US6348348-B1.  
 XX  
 PD 19-FEB-2002.  
 XX  
 PF 07-APR-1999; 99US-00287354.

QY	781	PQLWAAVGVSPHKGHLGKTNLCVEVADLVSLVHADTLPAPWHRACKQFLSGLDGEGLWS	840
Db	986	PQLWAAVGVSPHKGHLGKTNLCVEVADLVSLVHADAPLPAPWHRACKQFLSGLDGEGLWS	1045
QY	841	PGSQVSTVWVFRADQQRIRFLQWCPAGAGALEPGAGSCYLDAGLRRLREWGV	900
Db	1046	PGSQVSTVWVFRADQQRIRFLQWCPAGAGALEPGAGSCYLDAGLRRLREWGV	1105
QY	901	CWTLQAPGEAVLVPAGAPHQVQGLVSTVTHFLSPETSALSQQLCHQGSLLPDCHL	960
Db	1106	CWTLQAPGEAVLVPAGAPHQVQGLVSTVTHFLSPETSALSQQLCHQGSLLPDCHL	1165
QY	961	LYAQMDWAVFOAVKVAVGTLEAK	984
Db	1166	LYAQMDWAVFOAVKVAVGTLEAK	1189
RESULT 9			
ID	AAE19795	standard; protein; 1189 AA.	
XX	AAE19795;		
AC	AAE19795;		
XX	18-JUN-2002	(first entry)	
DT	18-JUN-2002	(first entry)	
XX	Human putative single zinc finger transcription factor protein.		
DE	Human putative single zinc finger transcription factor protein.		
XX	Human; Hairless protein; Hr protein; HR gene; dermatological condition;		
KW	hair loss; gene therapy; single zinc finger transcription factor protein.		
OS	Homo sapiens.		
XX	US6348348-B1.		
PN	19-FEB-2002.		
XX	07-APR-1999;	99US-00287354.	
PF	07-APR-1998;	98US-0080888P.	
XX	(CARN-) CARNEGIE INST WASHINGTON.		
PR	Thompson CC;		
XX	WPI; 2002-204622/26.		
DR	Novel expression construct, useful in the diagnosis and treatment of		
XX	dermatological conditions, such as hair loss, contains a Hairless gene		
PT	sequence.		
PS	Example; Fig 1; 48pp; English.		
XX	The invention relates to human Hairless (Hr) polypeptides and nucleic		
CC	acid molecules (HR) encoding such polypeptides. The invention also		
CC	relates to the Hairless expression constructs which may be used in		
CC	transcription assays. Sequences of the invention are used in diagnosis		
CC	and treatment of dermatological conditions such as hair loss. They are		
CC	also used in gene therapy. Polynucleotides of the invention can be used		
CC	as probes for the detection of hair loss. The present sequence is human		
CC	putative single zinc finger transcription factor protein (Hairless)		
XX	Sequence 1189 AA;		
QY	Query Match	98.7%;	Score 5290; DB 5; Length 1189;
Db	Best Local Similarity	99.0%;	Pred. No. 0;
QY	Matches 974; Conservative	1; Mismatches	9; Indels 0; Gaps 0;
Db	1	FYKDPISIPRLAKEPLAAEPGLFGLNSGHLQAGEAERPSLHQRDGEMGAGRQQNPPC	60
Db	206	FYKDPISIPRLAKEPLAAEPGLFGLNSGHLQAGEAERPSLHQRDGEMGAGRQQNPPC	265
QY	61	LFLGQPDTPWTSWPCPGLVHTLGNWAGPGDGNLGYQLGPPATPRCPSPFPVTQRG	120

Db	266	LFLGQPDTPWTSWPCPGLVHTLGNWAGPGDGNLGYQLGPPATPRCPSPFPVTQRG	325
QY	121	CCSSYPPTKGGDLGCGKCEGLEGGASGASBPSEVNKASGPACPPSHHTLKKTKWLT	180
Db	326	CCSSYPPTKGGDLGCGKCEGLEGGASGASBPSEVNKASGPACPPSHHTLKKTKWLT	385
QY	181	RHSEGFECPRGCEVEERPVARLRALKRAGSEVQGMGSPAPKRPDPFGPTAEQAGG	240
Db	386	RHSEGFECPRGCEVEERPVARLRALKRAGSEVQGMGSPAPKRPDPFGPTAEQAGG	445
QY	241	WQEVDRDTSIGNKDVDSGQHDQKPGDQOASIQDPLGLQDIPCLALPAKLAQCSQCAAG	300
Db	446	LQEVDRDTSIGNKDVDSGQHDQKPGDQOASIQDPLGLQDIPCLALPAKLAQCSQCAAG	505
QY	301	EGGGHACHSQVRRSPLGGELQOEEDTATNSSEEGPGSPDSRSLSTGLAKHLLSGLGR	360
Db	506	EGGGHACHSQVRRSPLGGELQOEEDTATNSSEEGPGSPDSRSLSTGLAKHLLSGLGR	565
QY	361	LCRLRRERELAWAQREGQGPVITGDSPIPRCCSRCHHGLFNTHWRCPSHRCLVAC	420
Db	566	LCRLRRERELAWAQREGQGPVITGDSPIPRCCSRCHHGLFNTHWRCPSHRCLVAC	625
QY	421	GRVAGTGRAREKAGFOEQSAEECTQAGHAACSLMLTQFVSSQALAEELSTAMHVVWKF	480
Db	626	GRVAGTGRAREKAGFOEQSAEECTQAGHAACSLMLTQFVSSQALAEELSTAMHVVWKF	685
QY	481	IRGHCPQADARVWAPGDAGQOKESTQKTPPTPQSCNGDTHRTKSIKEETPDSETPAE	540
Db	686	IRGHCPQADARVWAPGDAGQOKESTQKTPPTPQSCNGDTHRTKSIKEETPDSETPAE	745
QY	541	DRAGRGLPCPSLCELLASTAVKLCIGHRIHMAFAPVTPALPSPDDRIITNLDLSIAQVV	600
Db	746	DRAGRGLPCPSLCELLASTAVKLCIGHRIHMAFAPVTPALPSPDDRIITNLDLSIAQVV	805
QY	601	ERKIQEKALGPGLRAGPGLRKGILPLSPVRPLPPGALLMLQEPQCPRRGFHLFQEH	660
Db	806	ERKIQEKALGPGLRAGPGLRKGILPLSPVRPLPPGALLMLQEPQCPRRGFHLFQEH	865
QY	661	WFOGQPVLSGIQRTLQGNLWGTALGALGGQVQALSPGLPQPSLSLSTTTFWEGFSWPE	720
Db	866	WFOGQPVLSGIQRTLQGNLWGTALGALGGQVQALSPGLPQPSLSLSTTTFWEGFSWPE	925
QY	721	LRPKSDEGSVLLHRLAGDEDTSRVENLAASPLPEYCALHGKLNLSYLPGLALRPLE	780
Db	926	LRPKSDEGSVLLHRLAGDEDTSRVENLAASPLPEYCALHGKLNLSYLPGLALRPLE	985
QY	781	PQLWAAVGVSPHKGHLGKTNLCVEVADLVSLVHADTLPAPWHRACKQFLSLDGEGLWS	840
Db	986	PQLWAAVGVSPHKGHLGKTNLCVEVADLVSLVHADTLPAPWHRACKQFLSLDGEGLWS	1045
QY	841	PGSQVSTVWVFRADQQRIRFLQWCPAGAGALEPGAGSCYLDAGLRRLREWGV	900
Db	1046	PGSQVSTVWVFRADQQRIRFLQWCPAGAGALEPGAGSCYLDAGLRRLREWGV	1105
QY	901	CWTLQAPGEAVLVPAGAPHQVQGLVSTVTHFLSPETSALSQQLCHQGSLLPDCHL	960
Db	1106	CWTLQAPGEAVLVPAGAPHQVQGLVSTVTHFLSPETSALSQQLCHQGSLLPDCHL	1165
QY	961	LYAQMDWAVFOAVKVAVGTLEAK	984
Db	1166	LYAQMDWAVFOAVKVAVGTLEAK	1189
RESULT 10			
ID	ABU62541	standard; protein; 1189 AA.	
XX	ABU62541;		
XX	08-SEP-2003	(first entry)	
XX	Human Hairless (Hr) polypeptide #3.		

541 DRAGRLPCPSLCCELLASTAVKLCIGHERIHMAFAPVTPALPSDDRTNILDIIIAQV 600  
 746 DRAGRLPCPSLCCELLASTAVKLCIGHERIHMAFAPVTPALPSDDRTNILDIIIAQV 805  
 601 ERKIQEALGFCPLRAGPGLRKGGLPLSPVRPRLPPPGALLWLOPQPCPRRGFHLFOEH 660  
 806 ERKIQEALGFCPLRAGPGLRKGGLPLSPVRPRLPPPGALLWLOPQPCPRRGFHLFOEH 865  
 661 WRQGPVLVSGIORTLQNLWGTEALGALGGVQVQALSPGPPQPSLSGSTTTFWEGFSWPE 720  
 866 WRQGPVLVSGIORTLQNLWGTEALGALGGVQVQALSPGPPQPSLSGSTTTFWEGFSWPE 925  
 721 LRPKSDEGSVLLHRLALGDEDTSRVENLAASLPPEYCALHGKLNLSVLPGLALRPLE 780  
 926 LRPKSDEGSVLLHRLALGDEDTSRVENLAASLPPEYCALHGKLNLSVLPGLALRPLE 985  
 781 POLWAAVGVSPHRGHLGKTKNLCVEVADIVLSILVHADTLPALWRAQKDFLSGLDGEGLWS 840  
 986 POLWAAVGVSPHRGHLGKTKNLCVEVADIVLSILVHADTLPALWRAQKDFLSGLDGEGLWS 1045  
 841 PGSQVSTVWVFRAGDAQRIIRFLQWCPAGAGALEPGAGSCYLDAGLRRLREEWGVS 900  
 1046 PGSQVSTVWVFRAGDAQRIIRFLQWCPAGAGALEPGAGSCYLDAGLRRLREEWGVS 1105  
 901 CWTLLQAPGEAVLPAGAPHQVQGLVSTVSTOHLFSPETSALSAQLCHQGPSLPPDCHL 960  
 1106 CWTLLQAPGEAVLPAGAPHQVQGLVSTVSTOHLFSPETSALSAQLCHQGPSLPPDCHL 1165  
 961 LYAQMNAVFOAVKAVGTLOEAK 984  
 1166 LYAQMNAVFOAVKAVGTLOEAK 1189

## RESULT 8

AA15218  
 ID AA15218 standard; protein; 1189 AA.  
 XX  
 AC AA15218;  
 XX  
 DT 09-NOV-1999 (first entry)  
 XX  
 DE Human Hairless mutant amino acid sequence, Thr(1022)Ala.  
 XX  
 KW alopecia; congenital alopecia; congenital atrichia;  
 KW androgenetic alopecia; alopecia areata; alopecia universalis; mutant;  
 KW hair follicle.  
 XX  
 OS Homo sapiens.  
 XX  
 FH Key Location/Qualifiers  
 FT Misc-difference 1022  
 FT /label= T1022A  
 FT /note= "Wild-type Thr substituted by Ala"  
 XX  
 PN WO9938965-Al.  
 XX  
 XX 05-AUG-1999.  
 XX  
 XX 29-JAN-1999; 99WO-US002128.  
 XX  
 XX 29-JAN-1998; 98US-0073043P.  
 XX  
 XX (UYCO ) UNIV COLUMBIA NEW YORK.  
 XX  
 XX Christiano AM;  
 XX  
 XX WPI; 1999-479184/40.  
 XX  
 XX Human hairless gene and protein, useful for identifying modulators of  
 XX hair growth.  
 XX  
 XX Claim 8; Page; 127pp; English.  
 XX  
 XX

CC This is the amino acid sequence for the mutant human hairless protein,  
 CC which has Alanine substituted for Threonine at position 1022. The gene  
 CC was discovered by genotyping a Pakistani kindred (comprising of 4  
 CC affected males and 7 affected females) with an inherited form of  
 CC congenital alopecia universalis. The pedigree is strongly suggestive of  
 CC autosomal recessive inheritance. The invention provides methods and  
 CC sequences for the recombinant production of wild-type human hairless,  
 CC mutant human hairless and wild-type human whn (winged-helix-nude)  
 CC homologues, and animal models of hairlessness. Human hairless conditions  
 CC such as androgenetic alopecia (male pattern baldness), alopecia areata,  
 CC alopecia totalis, congenital alopecia universalis, congenital alopecia  
 CC and severe T-cell immunodeficiency can be treated with compounds  
 CC identified in the assays. The methods are also useful for identifying  
 CC compounds that can be used to inhibit hair growth  
 XX

SQ Sequence 1189 AA;

Query Match 99.4%; Score 5326; DB 2; Length 1189;  
 Best Local Similarity 99.6%; Pred. No. 0;  
 Matches 980; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 FYYKDPSTIPRLAKEPLAAABPGLFGLNSGGHLOPAGAEARPSLHQRDCEMAGCQQNCP 60  
 DB 206 FYYKDPSTIPRLAKEPLAAABPGLFGLNSGGHLOPAGAEARPSLHQRDCEMAGCQQNCP 265  
 QY 61 LFLGQPDTPVMTSWPACPPGLVHTLGNVWAGPGDNLGYQLGPPATPCPSPEPVTORG 120  
 DB 266 LFLGQPDTPVMTSWPACPPGLVHTLGNVWAGPGDNLGYQLGPPATPCPSPEPVTORG 325  
 QY 121 CCSSYPPTKGGDLGPOCKQBGLEGASGASEPSEEVNKASGAPRCPSSHHTKLKWTLT 180  
 DB 326 CCSSYPPTKGGDLGPOCKQBGLEGASGASEPSEEVNKASGAPRCPSSHHTKLKWTLT 385  
 QY 181 RHSEQFCPCRCPEVERPVARLRALKRAGSPVQGWGSPAPRPPDPPTGTAEQGAGG 240  
 DB 386 RHSEQFCPCRCPEVERPVARLRALKRAGSPVQGWGSPAPRPPDPPTGTAEQGAGG 445  
 QY 241 WQEVRTDSIGNKVDVSGQHDQKGPQDQASLOPGLQDIPCLALPAKLAQCCSCAQAG 300  
 DB 446 LQEVRTDSIGNKVDVSGQHDQKGPQDQASLOPGLQDIPCLALPAKLAQCCSCAQAG 505  
 QY 301 EGGGHACHSQOVRSPPLGGHLOQBEDTATNSSBEGSGSPDSRLSTGLAKHLISGLGDR 360  
 DB 506 EGGGHACHSQOVRSPPLGGHLOQBEDTATNSSBEGSGSPDSRLSTGLAKHLISGLGDR 565  
 QY 361 LCLLRERERALAWAOREGOGPAVTGSPGIPRCCSCHHGLFNTHWRCPCSHRLCVAC 420  
 DB 566 LCLLRERERALAWAOREGOGPAVTGSPGIPRCCSCHHGLFNTHWRCPCSHRLCVAC 625  
 QY 421 GRVAGTGRAREKAGFOQSABECTQEAHAACSLMLTQFVSSQALAEIISTAMHQQVVKFD 480  
 DB 626 GRVAGTGRAREKAGFOQSABECTQEAHAACSLMLTQFVSSQALAEIISTAMHQQVVKFD 685  
 QY 481 IRGHCPQOADARVWAPGDAGQOKESTQKTPPTQPSGNGTHRTKSIKEETPDSAETPAE 540  
 DB 686 IRGHCPQOADARVWAPGDAGQOKESTQKTPPTQPSGNGTHRTKSIKEETPDSAETPAE 745  
 QY 541 DRAGRGPLPCPSLCELLASTAVKLCIGHERIHMAFAPVTPALPSDDRTNILDIIIAQV 600  
 DB 746 DRAGRGPLPCPSLCELLASTAVKLCIGHERIHMAFAPVTPALPSDDRTNILDIIIAQV 805  
 QY 601 ERKIQEALGFCPLRAGPGLRKGGLPLSPVRPRLPPPGALLWLOPQPCPRRGFHLFOEH 660  
 DB 806 ERKIQEALGFCPLRAGPGLRKGGLPLSPVRPRLPPPGALLWLOPQPCPRRGFHLFOEH 865  
 QY 661 WRQGPVLVSGIORTLQNLWGTEALGALGGVQVQALSPGPPQPSLSGSTTTFWEGFSWPE 720  
 DB 866 WRQGPVLVSGIORTLQNLWGTEALGALGGVQVQALSPGPPQPSLSGSTTTFWEGFSWPE 925  
 QY 721 LRPKSDEGSVLLHRLALGDEDTSRVENLAASLPPEYCALHGKLNLSVLPGLALRPLE 780  
 DB 926 LRPKSDEGSVLLHRLALGDEDTSRVENLAASLPPEYCALHGKLNLSVLPGLALRPLE 985



QY 241 WQEVDTISGNKVDVSGQHDQKQPGDQASLQDPGLQDIPCLALPAKLAQCSCQAAG 300  
 Db 241 WQEVDTISGNKVDVSGQHDQKQPGDQASLQDPGLQDIPCLALPAKLAQCSCQAAG 300  
 QY 301 EGGGHACHSQOVRSPGLGELQOEEDTATNSSBEGPGSDSLSTGLAKHLLSLGDR 360  
 Db 301 EGGGHACHSQOVRSPGLGELQOEEDTATNSSBEGPGSDSLSTGLAKHLLSLGDR 360  
 QY 361 LCRLLRERERELAWAQREGQGPVATGSPGIPCCSRCHHGLFNTHWRCPCRSRLCVAC 420  
 Db 361 LCRLLRERERELAWAQREGQGPVATGSPGIPCCSRCHHGLFNTHWRCPCRSRLCVAC 420  
 QY 421 GRVAGTGRAREKAGFQSAEECTQEAAGHACSLMLTFVSSQALAEIETAMHQQVKFD 480  
 Db 421 GRVAGTGRAREKAGFQSAEECTQEAAGHACSLMLTFVSSQALAEIETAMHQQVKFD 480  
 QY 481 IRGHCPQADARVWAPGDAGQOKESTQKTPPTPOPCNGDTHRTKSIKEETPDSAETPAE 540  
 Db 481 IRGHCPQADARVWAPGDAGQOKESTQKTPPTPOPCNGDTHRTKSIKEETPDSAETPAE 540  
 QY 541 DRAGGBIPCSLCELLASTAVKICLGERIHMFAFVTPALPDDITNILDIIAQVV 600  
 Db 541 DRAGGBIPCSLCELLASTAVKICLGERIHMFAFVTPALPDDITNILDIIAQVV 600  
 QY 601 ERKIQEALGFLRAGFLRGKGLPLSPVRPRLPPGALLWLOEPQPCPRGHLFOEH 660  
 Db 601 ERKIQEALGFLRAGFLRGKGLPLSPVRPRLPPGALLWLOEPQPCPRGHLFOEH 660  
 QY 661 WRQGPVLVSGIQTLOGLNWGTALGALGQVQALSPGLPQPSSILGSTTFWGFWSPE 720  
 Db 661 WRQGPVLVSGIQTLOGLNWGTALGALGQVQALSPGLPQPSSILGSTTFWGFWSPE 720  
 QY 721 LRPKSDGSVLLLRALGDEBTSRVENLAASLPIPEYCALHKGKLNLSYLPDGLALRPLE 780  
 Db 721 LRPKSDGSVLLLRALGDEBTSRVENLAASLPIPEYCALHKGKLNLSYLPDGLALRPLE 780  
 QY 781 POLWAAVGVSPHRHGLGKNCVADVADLVSTLVHADTPLPAWHRAQKDFLSGLDGEGLWS 840  
 Db 781 POLWAAVGVSPHRHGLGKNCVADVADLVSTLVHADTPLPAWHRAQKDFLSGLDGEGLWS 840  
 QY 841 PGQVSTVWHVFRADAQRIRREFLQWCPAGAGALEPGAGSCYLDAGLRRLREEWGVS 900  
 Db 841 PGQVSTVWHVFRADAQRIRREFLQWCPAGAGALEPGAGSCYLDAGLRRLREEWGVS 900  
 QY 901 CWTLLQAPGEAVLPAGAPHQVGLVSTVSTQHFLSPETSAQAQLCHQGPSLPDCHL 960  
 Db 901 CWTLLQAPGEAVLPAGAPHQVGLVSTVSTQHFLSPETSAQAQLCHQGPSLPDCHL 960  
 QY 961 LYAQMNAVFOAVKAVGTTLQEA 984  
 Db 961 LYAQMNAVFOAVKAVGTTLQEA 984

RESULT 6  
 AAY15217  
 ID AAY15217 standard; protein; 1189 AA.  
 XX AC AAY15217;  
 XX DT 09-NOV-1999 (first entry)  
 XX DE Human Hairless wildtype protein amino acid sequence.  
 XX KW alopecia; congenital alopecia; congenital atrichia;  
 XX KW androgenetic alopecia; alopecia areata; alopecia universalis; wildtype;  
 XX KW hair follicle.  
 XX OS Homo sapiens.  
 XX EN W09938965-A1.  
 XX PD 05-AUG-1999.

XX 29-JAN-1999; 99WO-US002128.  
 XX 29-JAN-1998; 98US-0073043P.  
 XX (UYCO ) UNIV COLUMBIA NEW YORK.  
 XX Cristiano AM;  
 XX WPI; 1999-479184/40.  
 DR N-PSDB; AAZ06349.  
 XX Human hairless gene and protein, useful for identifying modulators of  
 hair growth.  
 XX Claim 21; Fig 6; 127pp; English.  
 XX This is the amino acid sequence for the wildtype human hairless protein.  
 CC The gene was discovered by genotyping a Pakistani kindred (comprising of  
 CC 4 affected males and 7 affected females) with an inherited form of  
 CC congenital alopecia universalis. The pedigree is strongly suggestive of  
 CC autosomal recessive inheritance. The invention provides methods and  
 CC sequences for the recombinant production of wild-type human hairless,  
 CC mutant human hairless and wild-type human whn (winged-helix-nude)  
 CC proteins, assays for screening for binding compounds, modulators and  
 CC homologues, and animal models of hairlessness. Human hairless conditions  
 CC such as androgenetic alopecia (male pattern baldness), alopecia areata,  
 CC alopecia totalis, congenital alopecia universalis, congenital alopecia  
 CC and severe T-cell immunodeficiency can be treated with compounds  
 CC identified in the assays. The methods are also useful for identifying  
 CC compounds that can be used to inhibit hair growth  
 XX Sequence 1189 AA;  
 QY Query Match 99.5%; Score 5331; DB 2; Length 1189;  
 Db Best Local Similarity 99.7%; Pred. No. 0;  
 QY Matches 981; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
 Db 1 FYKDPSIPRLAKEPLAAAEPLGLNSGGHQRAGEAERPSLHQRDGEAGAGQQNCP 60  
 206 FYKDPSIPRLAKEPLAAAEPLGLNSGGHQRAGEAERPSLHQRDGEAGAGQQNCP 265  
 QY 61 LFLGQDPTVPMKSWPACPGVLVHTLGNVWAGPGDNLGYQLGPATPCPSPEPVTVQ 120  
 Db 266 LFLGQDPTVPMKSWPACPGVLVHTLGNVWAGPGDNLGYQLGPATPCPSPEPVTVQ 325  
 QY 121 CCSSYPPTKGGDLGPCCKQEGLEGASGASEPSEEVNKASGPACPPSHHTKLKKTWLT 180  
 Db 326 CCSSYPPTKGGDLGPCCKQEGLEGASGASEPSEEVNKASGPACPPSHHTKLKKTWLT 385  
 QY 181 RHSEQFECPRGCEVEERPVARLALKRAGSPVQAGMSPAPKRPDPDPFGTAEQAGG 240  
 Db 386 RHSEQFECPRGCEVEERPVARLALKRAGSPVQAGMSPAPKRPDPDPFGTAEQAGG 445  
 QY 241 WQEVDTISGNKVDVSGQHDQKQPGDQASLQDPGLQDIPCLALPAKLAQCSCQAAG 300  
 Db 446 LQEVDTISGNKVDVSGQHDQKQPGDQASLQDPGLQDIPCLALPAKLAQCSCQAAG 505  
 QY 301 EGGGHACHSQOVRSPGLGELQOEEDTATNSSBEGPGSDSLSTGLAKHLLSLGDR 360  
 Db 506 EGGGHACHSQOVRSPGLGELQOEEDTATNSSBEGPGSDSLSTGLAKHLLSLGDR 565  
 QY 361 LCRLLRERERELAWAQREGQGPVATGSPGIPCCSRCHHGLFNTHWRCPCRSRLCVAC 420  
 Db 566 LCRLLRERERELAWAQREGQGPVATGSPGIPCCSRCHHGLFNTHWRCPCRSRLCVAC 625  
 QY 421 GRVAGTGRAREKAGFQSAEECTQEAAGHACSLMLTFVSSQALAEIETAMHQQVKFD 480  
 Db 626 GRVAGTGRAREKAGFQSAEECTQEAAGHACSLMLTFVSSQALAEIETAMHQQVKFD 685  
 QY 481 IRGHCPQADARVWAPGDAGQOKESTQKTPPTPOPCNGDTHRTKSIKEETPDPAETPAE 540  
 Db 686 IRGHCPQADARVWAPGDAGQOKESTQKTPPTPOPCNGDTHRTKSIKEETPDPAETPAE 745

QY 541 DRAGRLPCSLCELLASTAVKLCIGHERIHMAFAPVTPALPDDRIITNILDIIIAQV 600  
|||||  
Db 746 DRAGRLPCSLCELLASTAVKLCIGHERIHMAFAPVTPALPDDRIITNILDIIIAQV 805  
|||||  
QY 601 ERKIQEKALGPGRLAGPGLRGKGLPLSPVRPRLPPPGALLWLQFPQPCPRGRFHLFQSH 660  
|||||  
Db 806 ERKIQEKALGPGRLAGPGLRGKGLPLSPVRPRLPPPGALLWLQFPQPCPRGRFHLFQSH 865  
|||||  
QY 661 WRQGPVLVSGIORTLQGNLWTEALGALGGQVQALSPGPPQPSLSGTTTWEGFSWPE 720  
|||||  
Db 866 WRQGPVLVSGIORTLQGNLWTEALGALGGQVQALSPGPPQPSLSGTTTWEGFSWPE 925  
|||||  
QY 721 LRPKSDGSEVLLHRLAGDEDSRVENLAASPLPEYCALHGKLNLASVLPGLALRPLE 780  
|||||  
Db 926 LRPKSDGSEVLLHRLAGDEDSRVENLAASPLPEYCALHGKLNLASVLPGLALRPLE 985  
|||||  
QY 781 POLWAAAYGVSPPRHGLGTKNLCEVADVLSIIVHADTPLPAWHRAQKOFSLGLDGEGLWS 840  
|||||  
Db 986 POLWAAAYGVSPPRHGLGTKNLCEVADVLSIIVHADTPLPAWHRAQKOFSLGLDGEGLWS 1045  
|||||  
QY 841 PGSOYSTVWHVFRADOAQRIRFLOMCPAGAGALEPGAPGSCYLDAGLRRLREBWGS 900  
|||||  
Db 1046 PGSOYSTVWHVFRADOAQRIRFLOMCPAGAGALEPGAPGSCYLDAGLRRLREBWGS 1105  
|||||  
QY 901 CWTLLQAPGEAVLPAGAPHQVQGLVSTVSTQHLSPETSALSQALCHQGPSLPPDCHL 960  
|||||  
Db 1106 CWTLLQAPGEAVLPAGAPHQVQGLVSTVSTQHLSPETSALSQALCHQGPSLPPDCHL 1165  
|||||  
QY 961 LYAQMDFVQAVKXVAGTLOEAK 984  
|||||  
Db 1166 LYAQMDFVQAVKXVAGTLOEAK 1189  
|||||  
RESULT 5  
ABU62540  
ID ABU62540 standard; protein; 984 AA.  
XX AC ABU62540;  
XX DT 08-SEP-2003 (first entry)  
XX Human Hairless (Hr) polypeptide #1.  
DE Human; hairless; Hr; thyroid hormone receptor; hair growth; hair loss;  
KW family pedigree; radiation hybrid; somatic cell hybrid.  
XX OS Homo sapiens.  
FH Key Location/Qualifiers  
FT Misc-difference 50 /note= "Encoded by ATC"  
FT Misc-difference 140 /note= "Encoded by CAG"  
FT Misc-difference 384 /note= "Encoded by CTC"  
FT Misc-difference 446 /note= "Encoded by GAG"  
FT Misc-difference 453 /note= "Residue printed as Her"  
FT Misc-difference 467 /note= "Encoded by GAG"  
FT Misc-difference 529.530 /note= "Encoded by GAGGAG"  
FT Misc-difference 603 /note= "Encoded by AAC"  
FT Misc-difference 606 /note= "Encoded by GAG"  
FT Misc-difference 645 /note= "Encoded by GAG"  
FT Misc-difference 649 /note= "Residue printed as Gys"  
FT Misc-difference 659

FT Misc-difference 731 /note= "Encoded by GAG"  
FT /note= "Residue printed as Leo"  
FT Misc-difference 747 /note= "Encoded by AAG"  
XX US2003027300-A1.  
XX 06-FEB-2003.  
XX 21-DEC-2001; 2001US-00024369.  
XX 07-APR-1998; 98US-0080888F.  
XX 07-APR-1999; 99US-00287354.  
XX (CARN-) CARNEGIE INST WASHINGTON.  
XX Thompson CC;  
XX WPI; 2003-492034/58.  
XX N-PSDB; ACD26351.  
XX New Hairless polypeptide and polynucleotide, useful for identifying and  
PT detecting this genetic marker in family pedigrees or human-rodent somatic  
PT cell hybrids, or detecting interacting proteins that bind hairless gene  
PT or protein.  
XX Claim 1; Page 14-17; 34pp; English.  
PS The invention relates to a human hairless (HR) polynucleotide encoding  
XX the hairless (Hr) polypeptide. The invention also relates to an  
CC expression system comprising an expression construct which produces a  
CC polypeptide with hairless transcription factor activity, a reporter  
CC construct comprising a transcription regulatory region responsive to  
CC hairless transcription activity to regulate transcription of the reporter  
CC gene which is mediated by the transcription regulatory region and methods  
CC of screening for chemical agents which modulate hairless-mediated  
CC transcription, binding between hairless and thyroid hormone receptor or  
CC hairless activity. The polynucleotide is useful as a probe or primer to  
CC quantitate cognate RNA and DNA within cells, which can be subsequently  
CC used to correlate hair growth or loss with hairless expression or  
CC hairless-regulated transcription. The amino acid sequence of hairless  
CC antigen can be used for preparing specific binding molecules (e.g.  
CC polyclonal or monoclonal antibodies) for monitoring protein expression,  
CC for affinity purification and for functional studies. The human hairless  
CC polynucleotide, polypeptide or specific binding molecule may be used to  
CC identify and detect this genetic marker in family pedigrees, radiation  
CC hybrids or human-rodent somatic cell hybrids, and as an affinity tag to  
CC identify, isolate and detect interacting proteins that bind the hairless  
CC gene or protein. This sequence represents a human hairless polypeptide of  
CC the invention  
XX Sequence 984 AA;  
SQ Query Match 99.6%; Score 5334; DB 6; Length 984;  
Best Local Similarity 99.2%; Pred. No. 0;  
Matches 976; Conservative 8; Mismatches 0; Indels 0; Gaps 0;  
QY 1 FYYKPSIPRLAKEPLAAAEPLGLNLSGHLQAGEAERPSLHORDGEMGAGQONPCP 60  
|||||  
Db 1 FYYKPSIPRLAKEPLAAAEPLGLNLSGHLQAGEAERPSLHORDGEMGAGQONPCP 60  
|||||  
QY 61 LFLGQDTPVPTSWPACPPGLVHTLGNWAGDGNLGYQLGPPATPRCPSPPPVTQRG 120  
|||||  
Db 61 LFLGQDTPVPTSWPACPPGLVHTLGNWAGDGNLGYQLGPPATPRCPSPPPVTQRG 120  
|||||  
QY 121 CCSSVPTTKGDLGPCGKCEGLEGGASBPSEVNKASGPACPPSHHTLTKTTLT 180  
|||||  
Db 121 CCSSVPTTKGDLGPCGKCEGLEGGASBPSEVNKASGPACPPSHHTLTKTTLT 180  
|||||  
QY 181 RHSEQECPRGCPVEEERVARLRALKRAGSPEVQAGMSAPKRPDPPTGTAEOGAGG 240  
|||||  
Db 181 RHSEQECPRGCPVEEERVARLRALKRAGSPEVQAGMSAPKRPDPPTGTAEOGAGG 240  
|||||



QY 361 LCRLRRERREALAWAQREGQGPVATGDSGPGIPRCCSRCHHGLFNTHWPCPRCSHRLCVAC 420  
 Db |||||  
 QY 566 LCRLRRERREALAWAQREGQGPVATGDSGPGIPRCCSRCHHGLFNTHWPCPRCSHRLCVAC 625  
 Db |||||  
 QY 421 GRVAGTGAREKAGFQEQSABECTQEAAGAACSLMLTQFVSSQALAEELSTAMHGVVWVXF 480  
 Db |||||  
 QY 626 GRVAGTGAREKAGFQEQSABECTQEAAGAACSLMLTQFVSSQALAEELSTAMHGVVWVXF 685  
 Db |||||  
 QY 481 IRGHCPQADARVWAPGDAGQKQSTQKTPPTPQSCNGDTHRTKSIKEETPDSAETPAE 540  
 Db |||||  
 QY 686 IRGHCPQADARVWAPGDAGQKQSTQKTPPTPQSCNGDTHRTKSIKEETPDSAETPAE 745  
 Db |||||  
 QY 541 DRAGGPPCPCLCELLASTAVKLCGLHERTHMAFVTPALPSDDRTNILDSTIAQV 600  
 Db |||||  
 QY 746 DRAGGPPCPCLCELLASTAVKLCGLHERTHMAFVTPALPSDDRTNILDSTIAQV 805  
 Db |||||  
 QY 601 ERKIQEAKGFLRAGPGLRKGGLPLSPFVRPRLPPPGALLWLOEPQPCPRRGFHLFOEH 660  
 Db |||||  
 QY 806 ERKIQEAKGFLRAGPGLRKGGLPLSPFVRPRLPPPGALLWLOEPQPCPRRGFHLFOEH 865  
 Db |||||  
 QY 661 WRQGPVLVSGIORTLQNLGTEALGALGGVQVQALSPGPPQSSIGSTTFWEGFSWPE 720  
 Db |||||  
 QY 866 WRQGPVLVSGIORTLQNLGTEALGALGGVQVQALSPGPPQSSIGSTTFWEGFSWPE 925  
 Db |||||  
 QY 721 LRPKSDEGSVLLHRLGDEDTSRVENLAASLPLPEYCALHGKLNLASYLPPGLALRLE 780  
 Db |||||  
 QY 926 LRPKSDEGSVLLHRLGDEDTSRVENLAASLPLPEYCALHGKLNLASYLPPGLALRLE 985  
 Db |||||  
 QY 781 POLMAAGVSPHRGHLGKTNLCVEADVLSILVHADTLPAMHRAQKDFLSGLDGEGLWS 840  
 Db |||||  
 QY 986 POLMAAGVSPHRGHLGKTNLCVEADVLSILVHADTLPAMHRAQKDFLSGLDGEGLWS 1045  
 Db |||||  
 QY 841 PGSOVSTVWHVFRACDAQIRRFLOMVCPCAGALEPCAGSCYLDAGLRRLREWCVS 900  
 Db |||||  
 QY 1046 PGSOVSTVWHVFRACDAQIRRFLOMVCPCAGALEPCAGSCYLDAGLRRLREWCVS 1105  
 Db |||||  
 QY 901 CWTLLQAPGEAVLPAGAPHOVQGLVSTVYTHFLSPETSALSQALCHQGPSLPPDCHL 960  
 Db |||||  
 QY 1106 CWTLLQAPGEAVLPAGAPHOVQGLVSTVYTHFLSPETSALSQALCHQGPSLPPDCHL 1165  
 Db |||||  
 QY 961 LYAQMDWAVFOAVKAVGTQLEAK 984  
 Db |||||  
 QY 1166 LYAQMDWAVFOAVKAVGTQLEAK 1189  
 Db |||||

RESULT 4  
 ABU62542  
 ID ABU62542 standard; protein; 1189 AA.  
 XX AC ABU62542;  
 XX AC  
 XX DT 08-SEP-2003 (first entry)  
 XX DE Human Hairless (Hr) polypeptide #4.  
 XX KW Human; hairless; Hr; thyroid hormone receptor; hair growth; hair loss;  
 XX KW family pedigree; radiation hybrid; somatic cell hybrid.  
 XX OS Homo sapiens.  
 XX PN US2003027300-A1.  
 XX PD 06-FEB-2003.  
 XX PF 21-DEC-2001; 2001US-00024368.  
 XX PR 07-APR-1998; 98US-0080888P.  
 XX PR 07-APR-1999; 99US-00287354.  
 XX PA (CARN-) CARNegie INST WASHINGTON.  
 XX PI Thompson CC;  
 XX

WPI; 2003-492034/58.

New Hairless polypeptide and polynucleotide, useful for identifying and detecting this genetic marker in family pedigrees or human-rodent somatic cell hybrids, or detecting interacting proteins that bind hairless gene or protein.

Example; Fig 1; 34pp; English.

The invention relates to a human hairless (HR) polynucleotide encoding the hairless (Hr) polypeptide. The invention also relates to an expression system comprising an expression construct which produces a polypeptide with hairless transcription factor activity, a reporter construct comprising a transcription regulatory region responsive to hairless transcription activity to regulate transcription of the reporter gene which is mediated by the transcription regulatory region and methods of screening for chemical agents which modulate hairless-mediated transcription, binding between hairless and thyroid hormone receptor or hairless activity. The polynucleotide is useful as a probe or primer to quantitate cognate RNA and DNA within cells, which can be subsequently used to correlate hair growth or loss with hairless expression or hairless-regulated transcription. The amino acid sequence of hairless antigen can be used for preparing specific binding molecules (e.g. polyclonal or monoclonal antibodies) for monitoring protein expression, for affinity purification and for functional studies. The human hairless polynucleotide, polypeptide or specific binding molecule may be used to identify and detect this genetic marker in family pedigrees, radiation hybrids or human-rodent somatic cell hybrids, and as an affinity tag to identify, isolate and detect interacting proteins that bind the hairless gene or protein. This sequence represents a human hairless polypeptide of the invention

Sequence 1189 AA;

Query Match 99.7%; Score 5343; DB 6; Length 1189;

Best Local Similarity 99.8%; Pred. No. 0;

Matches 982; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 FYYKDPISIRLAKEPLAAAEPLGLNSGGHLPAGEAERPSLHQRDGEMGAGRQQPCP 60  
 Db |||||  
 QY 206 FYYKDPISIRLAKEPLAAAEPLGLNSGGHLPAGEAERPSLHQRDGEMGAGRQQPCP 265  
 Db |||||  
 QY 61 LFLGQPDTPVWTSWPACPPGLVHTLGNVWAGPDGNLGYQLGPPATPRCPSPPEVTVQ 120  
 Db |||||  
 QY 266 LFLGQPDTPVWTSWPACPPGLVHTLGNVWAGPDGNLGYQLGPPATPRCPSPPEVTVQ 325  
 Db |||||  
 QY 121 CCSSYPPTKGGDLGPGCKCGLEGGASGASEPEEVNKASGPRACPPSHHTKLKTKWLT 180  
 Db |||||  
 QY 326 CCSSYPPTKGGDLGPGCKCGLEGGASGASEPEEVNKASGPRACPPSHHTKLKTKWLT 385  
 Db |||||  
 QY 181 RHSEQFECPRGCPVEERPVARLRAKLAGSPVEQAGMSPPAPKRPDPFPFGTAEOGAGG 240  
 Db |||||  
 QY 386 RHSEQFECPRGCPVEERPVARLRAKLAGSPVEQAGMSPPAPKRPDPFPFGTAEOGAGG 445  
 Db |||||  
 QY 241 WQVRDTSIGNKVDSCQHDQKPGQDQASLQDPGLQDIPCLALPAKLAQCQCAQAG 300  
 Db |||||  
 QY 446 WQVRDTSIGNKVDSCQHDQKPGQDQASLQDPGLQDIPCLALPAKLAQCQCAQAG 505  
 Db |||||  
 QY 301 EGGHACHSQVRRSPGLGELQOEEDTATNSSSEEGPGSGPDSRLSTGLAKHLLSGLDGR 360  
 Db |||||  
 QY 506 EGGHACHSQVRRSPGLGELQOEEDTATNSSSEEGPGSGPDSRLSTGLAKHLLSGLDGR 565  
 Db |||||  
 QY 361 LCRLRRERREALAWAQREGQGPVATGDSGPGIPRCCSRCHHGLFNTHWPCPRCSHRLCVAC 420  
 Db |||||  
 QY 566 LCRLRRERREALAWAQREGQGPVATGDSGPGIPRCCSRCHHGLFNTHWPCPRCSHRLCVAC 625  
 Db |||||  
 QY 421 GRVAGTGAREKAGFQEQSABECTQEAAGAACSLMLTQFVSSQALAEELSTAMHGVVWVXF 480  
 Db |||||  
 QY 626 GRVAGTGAREKAGFQEQSABECTQEAAGAACSLMLTQFVSSQALAEELSTAMHGVVWVXF 685  
 Db |||||  
 QY 481 IRGHCPQADARVWAPGDAGQKQSTQKTPPTPQSCNGDTHRTKSIKEETPDPAETPAE 540  
 Db |||||  
 QY 686 IRGHCPQADARVWAPGDAGQKQSTQKTPPTPQSCNGDTHRTKSIKEETPDPAETPAE 745  
 Db |||||

Best Local Similarity 99.9%; Pred. No. 0;		Matches 983; Conservative 0; Mismatches 1; Indels 0; Gaps 0;	
QY	1	FYYKDPSPRLAKEPLAAAEPLFGLNSGGHLQORAGEAERPSLHORDGEMGAGROONPCP	60
Db	1	FYYKDPSPRLAKEPLAAAEPLFGLNSGGHLQORAGEAERPSLHORDGEMGAGROONPCP	60
QY	61	LFLGQPDTPVWTSWPACPPGLVHTLGNWAGDGNLGYQLGPPATPRCPSPPEPVTORG	120
Db	61	LFLGQPDTPVWTSWPACPPGLVHTLGNWAGDGNLGYQLGPPATPRCPSPPEPVTORG	120
QY	121	CCSSYPPTKGGDLGPCGCKQEGLEGASGASPESEVNKASGPRACPPSHHTKLKKTWLT	180
Db	121	CCSSYPPTKGGDLGPCGCKQEGLEGASGASPESEVNKASGPRACPPSHHTKLKKTWLT	180
QY	181	RHSEQFECPRGCPVEEERPVARLRLKRAKSGSEVQAMGSPAPKRPDPFPGTAEGGAGG	240
Db	181	RHSEQFECPRGCPVEEERPVARLRLKRAKSGSEVQAMGSPAPKRPDPFPGTAEGGAGG	240
QY	241	WQEVRTSTGNKDVDSGQHDQKGPQDGQASLQDPGLQDIPCLALPAKLAQCQSCAAAG	300
Db	241	WQEVRTSTGNKDVDSGQHDQKGPQDGQASLQDPGLQDIPCLALPAKLAQCQSCAAAG	300
QY	301	EGGHACHSQVRRSPPLGGELQOEEDTATNSSSEEGPGSPDSRLSTGLAKHLLSLGDR	360
Db	301	EGGHACHSQVRRSPPLGGELQOEEDTATNSSSEEGPGSPDSRLSTGLAKHLLSLGDR	360
QY	361	LCRLRRERREALAWAREGQGAFTGDSPIPRCCSRCHHGFNTHWRCPRSHRLCVAC	420
Db	361	LCRLRRERREALAWAREGQGAFTGDSPIPRCCSRCHHGFNTHWRCPRSHRLCVAC	420
QY	421	GRVAGTGRAREKAGFQESAEECTQAGHAAACSLMTQFVSSQAALSTAMHQVWVKED	480
Db	421	GRVAGTGRAREKAGFQESAEECTQAGHAAACSLMTQFVSSQAALSTAMHQVWVKED	480
QY	481	IRGHCPQADARWAPGDAQQKESQTKPTTPQSPSCNGDTHRTKSIKEETPDSAEPTAE	540
Db	481	IRGHCPQADARWAPGDAQQKESQTKPTTPQSPSCNGDTHRTKSIKEETPDSAEPTAE	540
QY	541	DRAGGELPCPSLCHELLASTAVKLCIGHERIHWAPFVTPALPSDRIITNILDSTIAQVV	600
Db	541	DRAGGELPCPSLCHELLASTAVKLCIGHERIHWAPFVTPALPSDRIITNILDSTIAQVV	600
QY	601	ERKIQEKALGCLRAGPGLRKLGLPLSPVRPRLPPGALLMLQEPQPCPRRGFHLFQSH	660
Db	601	ERKIQEKALGCLRAGPGLRKLGLPLSPVRPRLPPGALLMLQEPQPCPRRGFHLFQSH	660
QY	661	WRQGPVLVSGIQTTLQGNLWTEALGALGGQVQALSPGLPQPSPSLGTTTWEGFSWPE	720
Db	661	WRQGPVLVSGIQTTLQGNLWTEALGALGGQVQALSPGLPQPSPSLGTTTWEGFSWPE	720
QY	721	LRPKDEGSLVLLHRLAGDEDTSRVENTLAASPLPEYCALHGKINLASVLPGLALRPLE	780
Db	721	LRPKDEGSLVLLHRLAGDEDTSRVENTLAASPLPEYCALHGKINLASVLPGLALRPLE	780
QY	781	POLMAAYGVSPHRHGLTKNLCEVADLVSLVHADTPLPAWHRAQKDFLSGLDGEGLWS	840
Db	781	POLMAAYGVSPHRHGLTKNLCEVADLVSLVHADTPLPAWHRAQKDFLSGLDGEGLWS	840
QY	841	PGSQVSTVWVFPRAQDAQIRPLQVPCPAGAGALEPGAPGSCYLDAGLRRLRREWGV	900
Db	841	PGSQVSTVWVFPRAQDAQIRPLQVPCPAGAGALEPGAPGSCYLDAGLRRLRREWGV	900
QY	901	CWTLLOAPCEAVLPAGAPHOVQGLVSTVSTQHFLSPETSALSQILCHQGSPSLPDDCHL	960
Db	901	CWTLLOAPCEAVLPAGAPHOVQGLVSTVSTQHFLSPETSALSQILCHQGSPSLPDDCHL	960
QY	961	LXAQMDWAVFQAVKVAVGTLOEAK 984	
Db	961	LXAQMDWAVFQAVKVAVGTLOEAK 984	

CC acid molecules (HR) encoding such polypeptides. The invention also  
 CC relates to the Hairless expression constructs which may be used in  
 CC transcription assays. Sequences of the invention are used in diagnosis  
 CC and treatment of dermatological conditions such as hair loss. They are  
 CC also used in gene therapy. Polynucleotides of the invention can be used  
 CC as probes for the detection of hair loss. The present sequence is human  
 CC Hairless protein (Hr)  
 XX  
 SQ Sequence 984 AA;

Query Match 100.0%; Score 5358; DB 5; Length 984;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 984; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FYYKDPSTPRLAKEPLAAAEFGFLGNSGGHLLORAGEAERPSLHORDGEMGAGROONPCP 60  
 DB 1 FYYKDPSTPRLAKEPLAAAEFGFLGNSGGHLLORAGEAERPSLHORDGEMGAGROONPCP 60

QY 61 LFLGQDTPVWTSWPACPPGLVHTLGNVWAGDGNLGYQLGPPATPCPSPEPVTORG 120  
 DB 61 LFLGQDTPVWTSWPACPPGLVHTLGNVWAGDGNLGYQLGPPATPCPSPEPVTORG 120

QY 121 CCSSYPPTKGGDLGFCGKQCELEGASGASEPSEEVNKAAGPRACPPSHHTKLKKTWLT 180  
 DB 121 CCSSYPPTKGGDLGFCGKQCELEGASGASEPSEEVNKAAGPRACPPSHHTKLKKTWLT 180

QY 181 RSHQFECFCGCEVEERPVARLRLKAGSPFVQAGSGAPKRPDPFGTAEQAGG 240  
 DB 181 RSHQFECFCGCEVEERPVARLRLKAGSPFVQAGSGAPKRPDPFGTAEQAGG 240

QY 241 WQEVDTSIGNKVDSDGQHQKPDQCAQLQDIPCLALPAKLAOCQCAQAG 300  
 DB 241 WQEVDTSIGNKVDSDGQHQKPDQCAQLQDIPCLALPAKLAOCQCAQAG 300

QY 301 EGGHACHSQOVRSPPLGGELQOEEDTATNSSBEGPGSGPDSRLSTGLAKHLLSLGDR 360  
 DB 301 EGGHACHSQOVRSPPLGGELQOEEDTATNSSBEGPGSGPDSRLSTGLAKHLLSLGDR 360

QY 361 LCRLLRERERALAWOREGQPAVTDSPGTPRCSRCHGLFNTHWRCPCSHRLCVAC 420  
 DB 361 LCRLLRERERALAWOREGQPAVTDSPGTPRCSRCHGLFNTHWRCPCSHRLCVAC 420

QY 421 GRVAGTGRAREKAGFQESAECTQEAHGAACSLMTQFVSSQALAEI STAMHVVWKFED 480  
 DB 421 GRVAGTGRAREKAGFQESAECTQEAHGAACSLMTQFVSSQALAEI STAMHVVWKFED 480

QY 481 IRGHCPCQADARVWAPGDAGQKQESTQKTPPTQPSCNGDTHRTKSIKEETPDSAETPAE 540  
 DB 481 IRGHCPCQADARVWAPGDAGQKQESTQKTPPTQPSCNGDTHRTKSIKEETPDSAETPAE 540

QY 541 DRAGRGPLPCSLCELLASTAVKLCIGHERIHMFAFAPVTPALPDDDRITNILDIIIAQVV 600  
 DB 541 DRAGRGPLPCSLCELLASTAVKLCIGHERIHMFAFAPVTPALPDDDRITNILDIIIAQVV 600

QY 601 ERKIQKALGFLRAGPLRKGLGLPSVPRPLPPGALLWLOEPCPRGRGPHLQEH 660  
 DB 601 ERKIQKALGFLRAGPLRKGLGLPSVPRPLPPGALLWLOEPCPRGRGPHLQEH 660

QY 661 WRQCPVLVSGIQTLQNLWGTEALGALGQVQALSPGPPQPSLSTFTWEGFSWPE 720  
 DB 661 WRQCPVLVSGIQTLQNLWGTEALGALGQVQALSPGPPQPSLSTFTWEGFSWPE 720

QY 721 LRPKSDGVSLLHRLAGDEDTSEVENLAASLPEYCALHGKLNLAAYLPGLALRPLE 780  
 DB 721 LRPKSDGVSLLHRLAGDEDTSEVENLAASLPEYCALHGKLNLAAYLPGLALRPLE 780

QY 781 POLWAAVGVSPHRLGTLKNCVAVDLVSLVHADTPLPAWHRAQKDFLSGLDGEGLWS 840  
 DB 781 POLWAAVGVSPHRLGTLKNCVAVDLVSLVHADTPLPAWHRAQKDFLSGLDGEGLWS 840

QY 841 PGQSVSTVWVFRQAQRIIRRFLOMVCPCAGALEPFCAGSCYLDAGLRRLREEWGVS 900  
 DB 841 PGQSVSTVWVFRQAQRIIRRFLOMVCPCAGALEPFCAGSCYLDAGLRRLREEWGVS 900

QY 901 CWTLLQAPGEAVLPAGAPHQVQGLVSTVYTHFLSPETSALSALCHQGPSLPDCHL 960  
 DB 901 CWTLLQAPGEAVLPAGAPHQVQGLVSTVYTHFLSPETSALSALCHQGPSLPDCHL 960

QY 961 LYAQMDWAVFOAVKAVGTLQEA 984  
 DB 961 LYAQMDWAVFOAVKAVGTLQEA 984

RESULT 2  
 ABU62545  
 ID ABU62545 standard; protein; 984 AA.  
 XX  
 AC ABU62545;  
 XX  
 DT 08-SEP-2003 (first entry)  
 XX  
 DE Human Hairless (Hr) polypeptide #2.  
 XX  
 KW Human; hairless; Hr; thyroid hormone receptor; hair growth; hair loss;  
 KW family pedigree; radiation hybrid; somatic cell hybrid.  
 XX  
 OS Homo sapiens.  
 XX  
 FN US2003027300-A1.  
 XX  
 PD 06-FEB-2003.  
 XX  
 PF 21-DEC-2001; 2001US-00024368.  
 XX  
 ER 07-APR-1998; 98US-0080888P.  
 PR 07-APR-1999; 99US-00287354.  
 XX  
 PA (CARN-) CARNEGIE INST WASHINGTON.  
 XX  
 PI Thompson CC;  
 XX  
 DR WPI; 2003-492034/58.  
 XX  
 PT New Hairless polypeptide and polynucleotide, useful for identifying and  
 PT detecting this genetic marker in family pedigrees or human-rodent somatic  
 PT cell hybrids, or detecting interacting proteins that bind hairless gene  
 PT or protein.  
 XX

Example; Fig 1; 34pp; English.

The invention relates to a human hairless (HR) polynucleotide encoding  
 the hairless (Hr) polypeptide. The invention also relates to an  
 expression system comprising an expression construct which produces a  
 polypeptide with hairless transcription factor activity, a reporter  
 construct comprising a transcription regulatory region responsive to  
 hairless transcription activity to regulate transcription of the reporter  
 gene which is mediated by the transcription regulatory region and methods  
 of screening for chemical agents which modulate hairless-mediated  
 transcription, binding between hairless and thyroid hormone receptor or  
 hairless activity. The polynucleotide is useful as a probe or primer to  
 quantitate cognate RNA and DNA within cells, which can be subsequently  
 used to correlate hair growth or loss with hairless expression or  
 hairless-regulated transcription. The amino acid sequence of hairless  
 antigen can be used for preparing specific binding molecules (e.g.  
 polyclonal or monoclonal antibodies) for monitoring protein expression,  
 for affinity purification and for functional studies. The human hairless  
 polynucleotide, polypeptide or specific binding molecule may be used to  
 identify and detect this genetic marker in family pedigrees, radiation  
 hybrids or human-rodent somatic cell hybrids, and as an affinity tag to  
 identify, isolate and detect interacting proteins that bind the hairless  
 gene or protein. This sequence represents a human hairless polypeptide of  
 the invention

Sequence 984 AA;

Query Match 99.9%; Score 5353; DB 6; Length 984;

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 29, 2004, 10:19:04 ; Search time 65 Seconds  
(without alignments)  
4277.335 Million cell updates/sec

Title: US-10-024-368-2

Perfect score: 5358

Sequence: 1 FYYKDPSPRLAKEPLAAAE.....MDWAVQAVKVAAGTLOBAK 984

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A\_Geneseq 25Jan04:\*

- 1: Geneseqp1980s:\*
- 2: Geneseqp1990s:\*
- 3: Geneseqp2000s:\*
- 4: Geneseqp2001s:\*
- 5: Geneseqp2002s:\*
- 6: Geneseqp2003as:\*
- 7: Geneseqp2003bs:\*
- 8: Geneseqp2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	5358	100.0	984	5	Aae19794	Aae19794 Human Hal
2	5353	99.9	984	6	ABU62545	ABU62545 Human Hal
3	5343	99.7	1189	5	Aae19796	Aae19796 Human Hal
4	5343	99.7	1189	6	ABU62542	ABU62542 Human Hal
5	5334	99.6	984	6	ABU62540	ABU62540 Human Hal
6	5331	99.5	1189	2	AAV15217	AAV15217 Human Hal
7	5331	99.5	1189	6	ABG72775	ABG72775 Human Hal
8	5326	99.4	1189	2	AAV15218	AAV15218 Human Hal
9	5290	98.7	1189	5	Aae19795	Aae19795 Human put
10	5290	98.7	1189	6	ABU62541	ABU62541 Human Hal
11	4102.5	76.6	1182	5	Aae19798	Aae19798 Mouse Hal
12	4102.5	76.6	1182	6	ABU62544	ABU62544 Mouse Hal
13	4067	75.9	1207	5	Aae19797	Aae19797 Rat Hair1
14	4067	75.9	1207	6	ABU62543	ABU62543 Rat Hair1
15	965	18.0	179	6	ABU70953	ABU70953 Human adi
16	777.5	14.5	1266	4	AAM40172	AAM40172 Human nuc
17	777.5	14.5	2055	8	ADE40509	ADE40509 Human pol
18	711.5	13.3	1145	3	AAB42533	AAB42533 Human ORF
19	707	13.2	1337	4	AAM78460	AAM78460 Human pro
20	707	13.2	1338	4	AAM79444	AAM79444 Human pro
21	703	13.1	1236	7	ADD18797	ADD18797 Human dis
22	696.5	13.0	1212	5	ABP65070	ABP65070 Hypoxia-i
23	696.5	13.0	1213	7	ADD46413	ADD46413 Human Pro
24	694.5	13.0	1214	7	ADD46411	ADD46411 Rat Prote
25	574	10.7	671	5	ABP41598	ABP41598 Human ova

ALIGNMENTS

RESULT 1  
AAE19794  
ID AAE19794 standard; protein; 984 AA.  
XX  
AC AAE19794;  
XX  
DT 18-JUN-2002 (first entry)  
XX  
DE Human Hairless protein (Hr) #1.  
XX  
KW Human; Hairless protein; Hr protein; HR gene; dermatological condition;  
KW hair loss; gene therapy.  
XX  
OS Homo sapiens.  
XX  
FH Key  
FT Misc-difference 284  
FT /note= "Encoded by CTT of the sequence shown as SEQ ID  
FT NO:1 (AAD31072) in Column 35-38 of the specification"  
FT Misc-difference 978  
FT /note= "Encoded by CGG of the sequence shown as SEQ ID  
FT NO:1 (AAD31077) in Column 25-32 of the specification"  
FT Misc-difference 982  
FT /note= "Encoded by CAG of the sequence shown as SEQ ID  
FT NO:1 (AAD31077) in Column 25-32 of the specification"

US6348348-B1.

19-FEB-2002.

07-APR-1999; 99US-00287354.

07-APR-1998; 98US-0080888P.

(CARN-) CARNEGIE INST WASHINGTON.

Thompson CC;

WPI; 2002-204622/26.

N-PSDB; AAD31072, AAD31077.

Novel expression construct, useful in the diagnosis and treatment of dermatological conditions, such as hair loss, contains a Hairless gene sequence.

Claim 1b; Fig 1; 48pp; English.

The invention relates to human Hairless (Hr) polypeptides and nucleic

Abb63729 Drosophil  
Aam1958 Human pol  
Ada55130 Human pol  
Abg22993 Novel hum  
Abg93938 Human pol  
Aab40638 Human ORF  
Aam40312 Human pol  
Adc39252 Novel hum  
Aam79494 Human pro  
Ade09425 Novel pro  
Ade73110 Human cel  
Aam78510 Human pro  
Abg03533 Novel hum  
Abp76682 Streptom  
Adp76682 Human nov  
Adc33247 Human nov  
Adc86801 Human GPC  
Abr57561 Human MC2  
Abg23029 Novel hum  
Abg15105 Novel hum



Db 5220 CCAACACACAGTGGCCACACAGCTCACACCTGTCTCTCAGGCTGGCACTCTCCCCAC 5279  
Qy 3181 CCTGTGCCCTTTT 3193  
Db 5280 CCTGTGCCCTTTT 5292

RESULT 3  
US-10-414-692-34  
; Sequence 34, Application US/10414692  
; Publication No. US20030228607A1  
; GENERAL INFORMATION:  
; APPLICANT: X-Ceptor Therapeutics, Inc.  
; TITLE OF INVENTION: Screening method and modulators having an improved therapeutic  
; FILE OF INVENTION: profile  
; FILE REFERENCE: 8012-002-US  
; CURRENT APPLICATION NUMBER: US/10/414,692  
; CURRENT FILING DATE: 2003-04-14  
; PRIOR APPLICATION NUMBER: 60/372,650  
; PRIOR FILING DATE: 2002-04-15  
; NUMBER OF SEQ ID NOS: 86  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 34  
; LENGTH: 3570  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-414-692-34

Query Match 92.0%; Score 2945.4; DB 16; Length 3570;  
Best Local Similarity 99.8%; Pred. No. 0;  
Matches 2949; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 1 TTTTACTACAAGATCCGAGCATTCACAGTTGGGAAAGAGCCCTTGGCAGCTGGGAA 60  
Db 616 TTTTACTACAAGATCCGAGCATTCACAGTTGGGAAAGAGCCCTTGGCAGCTGGGAA 675

Qy 61 CTGTGGTCTTGGCTTAACTCTGTGGGACCTCTGAGAGCCCGGGAGGCGGAACGC 120  
Db 676 CTGTGGTCTTGGCTTAACTCTGTGGGACCTCTGAGAGCCCGGGAGGCGGAACGC 735

Qy 121 CTTTCACTGACACAGAGGATGGAGATGGAGCTGGCGGAGCAGAAATCTTGGCCG 180  
Db 736 CTTTCACTGACACAGAGGATGGAGATGGAGCTGGCGGAGCAGAAATCTTGGCCG 795

Qy 181 CTCTTCTGGGGAGCAGACACTGTGCTTGAACCTCTGGGCGGCTTGTCTCCCGCAGGC 240  
Db 796 CTCTTCTGGGGAGCAGACACTGTGCTTGAACCTCTGGGCGGCTTGTCTCCCGCAGGC 855

Qy 241 CTTTGTATCACTCTTGGCAACGCTTGGGCTGGGCGAGGCGATGGGAACCTTGGGTACAG 300  
Db 856 CTTTGTATCACTCTTGGCAACGCTTGGGCTGGGCGAGGCGATGGGAACCTTGGGTACAG 915

Qy 301 CTGGGGCCACAGCAACACCAAGGTGCGCTCTCTGAGCGCGCTGTCAACCGAGGCG 360  
Db 916 CTGGGGCCACAGCAACACCAAGGTGCGCTCTCTGAGCGCGCTGTCAACCGAGGCG 975

Qy 361 TGTGTGTATCTTACCAACCACTAAAGGTGGGATCTTGGCCCTTGTGGGAGTGCAG 420  
Db 976 TGTGTGTATCTTACCAACCACTAAAGGTGGGAGTCTTGGCCCTTGTGGGAGTGCAG 1035

Qy 421 GAGGCGCTGAGGGGGTGGCAGTGGAGCCAGCGAACCCAGCGAGGAAGTGAACAAGGCC 480  
Db 1036 GAGGCGCTGAGGGGGTGGCAGTGGAGCCAGCGAACCCAGCGAGGAAGTGAACAAGGCC 1095

Qy 481 TCTGGCCCCAGGGGCTGTCCCCCAGCCACCAACCAAGTGAAGAGATGCTCTACA 540  
Db 1096 TCTGGCCCCAGGGGCTGTCCCCCAGCCACCAACCAAGTGAAGAGATGCTCTACA 1155

Qy 541 CGGCACCTGAGAGAGTTGAATGTCCAGCGGCTGCGGCTGAGGTGAGGAGGCGGTT 600  
Db 1156 CGGCACCTGAGAGAGTTGAATGTCCAGCGGCTGCGGCTGAGGTGAGGAGGCGGTT 1215

Qy 601 GCTCGGCTCCGGGCTCTCAAAAGGCGCAGCAGCCCGAGGTTCCAGGGAGCAATGGSCAGT 660  
Db 1216 GCTCGGCTCCGGGCTCTCAAAAGGCGCAGCAGCCCGAGGTTCCAGGGAGCAATGGSCAGT 1275

Qy 661 CCAGCCCCAAAGCGGCCACCGGACCTTTCCAGGCACTGACAGAAACAGGGGGCTGGGGGT 720  
Db 1276 CCAGCCCCAAAGCGGCCACCGGACCTTTCCAGGCACTGACAGAAACAGGGGGCTGGGGGT 1335

Qy 721 TGGCAGGAGGTTTCGGGACACATCGATAGGAAACAAGGATGTGGACTTCGGGACAGCATGAT 780  
Db 1336 TGGCAGGAGGTTTCGGGACACATCGATAGGAAACAAGGATGTGGACTTCGGGACAGCATGAT 1395

Qy 781 GAGCAAGAAAGGACCCCAAGATGGCCAGCCACTCTCAGGACCCCGGACTTCAGGACATA 840  
Db 1396 GAGCAAGAAAGGACCCCAAGATGGCCAGCCACTCTCAGGACCCCGGACTTCAGGACATA 1455

Qy 841 CCATCGCTGCTTCTCCCTGCAAACTGGCTCAATGCCAAAGTTGTGCCCAAGGAGCTGGA 900  
Db 1456 CCATCGCTGCTTCTCCCTGCAAACTGGCTCAATGCCAAAGTTGTGCCCAAGGAGCTGGA 1515

Qy 901 GAGGAGGAGGCGCAGCCTTCCACTCTCAGCAAGTGGAGATTCGCTCTGGGAGGGAG 960  
Db 1516 GAGGAGGAGGCGCAGCCTTCCACTCTCAGCAAGTGGAGATTCGCTCTGGGAGGGAG 1575

Qy 961 CTGACAGGAGGAGGACACAGCCACTCTCAGCTCTGAGGAGGCGCCAGGCTCCGGC 1020  
Db 1576 CTGACAGGAGGAGGACACAGCCACTCTCAGCTCTGAGGAGGCGCCAGGCTCCGGC 1635

Qy 1021 CTTGACAGCGGCTCAGACACAGCCCTCGCAAGCACTGCTCAGTGGTTTGGGGAGCCGA 1080  
Db 1636 CTTGACAGCGGCTCAGACACAGCCCTCGCAAGCACTGCTCAGTGGTTTGGGGAGCCGA 1695

Qy 1081 CTGTGCGGCTCTGTCGGAGGAGCGGAGGCGCTGGCTTGGGCGCCAGCGGAGGAGCCAA 1140  
Db 1696 CTGTGCGGCTCTGTCGGAGGAGCGGAGGCGCTGGCTTGGGCGCCAGCGGAGGAGCCAA 1755

Qy 1141 GGGCCAGCGCTGACAGGGGACAGCCAGGATTCACAGCTCTGACAGCGTTCGCCACCAT 1200  
Db 1756 GGGCCAGCGCTGACAGGGGACAGCCAGGATTCACAGCTCTGACAGCGTTCGCCACCAT 1815

Qy 1201 GGNACTCTTCAACACCACTGGCGATGTCCCGCTGACAGCCAGCGGCTGTGTGGCCTGT 1260  
Db 1816 GGNACTCTTCAACACCACTGGCGATGTCCCGCTGACAGCCAGCGGCTGTGTGGCCTGT 1875

Qy 1261 GGTCTGTGTGAGGCACTGGGCGGGCCAGGAGAAAGAGCGCTTTCAGGAGGAGTCCGCG 1320  
Db 1876 GGTCTGTGTGAGGCACTGGGCGGGCCAGGAGAAAGAGCGCTTTCAGGAGGAGTCCGCG 1935

Qy 1321 GAGGAGTGCAGCAGAGGCGGGGACGCTGCTGCTGTCCCTGATGCTGACCCAGTTTGTGTC 1380  
Db 1936 GAGGAGTGCAGCAGAGGCGGGGACGCTGCTGCTGTCCCTGATGCTGACCCAGTTTGTGTC 1995

Qy 1381 TCCAGCCAGGCTTTGGCAGAGCTGAGCACTGCAATGACCAAGGCTCTGGGTCAAGTTTGTAT 1440  
Db 1996 TCCAGCCAGGCTTTGGCAGAGCTGAGCACTGCAATGACCAAGGCTCTGGGTCAAGTTTGTAT 2055

Qy 1441 ATCCGGGGGCACTGCCCCCTCCCAAGTGTATCCCGGATATGGGCCCCCGGGGATGACGGC 1500  
Db 2056 ATCCGGGGGCACTGCCCCCTCCCAAGTGTATCCCGGATATGGGCCCCCGGGGATGACGGC 2115

Qy 1501 CAGCAGAGGAATCAACACAGAAACCGCCCACTCCACAACTTCTGCAATGGCGAC 1560  
Db 2116 CAGCAGAGGAATCAACACAGAAACCGCCCACTCCACAACTTCTGCAATGGCGAC 2175

Qy 1561 ACCCAGAGGACCAAGAGCATCAAGAGGAGAGACCCCGATTTCCGCTGAGACCCAGCAGAG 1620  
Db 2176 ACCCAGAGGACCAAGAGCATCAAGAGGAGAGACCCCGATTTCCGCTGAGACCCAGCAGAG 2235

Qy 1621 GACCGTCTGGCGAGGCGCCCTGCTTGTCTCTCTGCGAACTGCTGGCTTCTTACC 1680  
Db 2236 GACCGTCTGGCGAGGCGCCCTGCTTGTCTCTCTGCGAACTGCTGGCTTCTTACC 2295

Qy 1681 GCGGTCAAACTCTGCTTTGGGCCATCAGCATGGGCTTTCGCCCCCGTCACTCCG 1740





Db	2341	 CCCCAGCTCTGGG	CAGCCTTAGTGTGAGCCCGCACCGGGGACCTTGGGGACCAAGAAC	2400
Qy	2401	CTCTGTGTGAGGTGGCCGACCTGGTCAGCATCTCTGGTGATCGCGACACACACCTGCCT	2460	
Db	2401	CTCTGTGTGAGGTGGCCGACCTGGTCAGCATCTCTGGTGATCGCGACACACACCTGCCT	2460	
Qy	2461	GCCTGGCACCGGGCACAAAGAATCTCTTTCAGGCCTGGACGGGAGGGGCTCTGGTCT	2520	
Db	2461	GCCTGGCACCGGGCACAAAGAATCTCTTTCAGGCCTGGACGGGAGGGGCTCTGGTCT	2520	
Qy	2521	CCGGGACGACGCTCAGCACTGTGTGCACGTGTTCCGGGCA CAGGACGCCACAGCGATC	2580	
Db	2521	CCGGGACGACGCTCAGCACTGTGTGCACGTGTTCCGGGCA CAGGACGCCACAGCGATC	2580	
Qy	2581	CGCCGCTTTCTCCAGATGGTGTGCCCGCCGGGGCAGGCGCCCTTGAGCCTGGCGGCCCA	2640	
Db	2581	CGCCGCTTTCTCCAGATGGTGTGCCCGCCGGGGCAGGCGCCCTTGAGCCTGGCGGCCCA	2640	
Qy	2641	GGCAGCTGCTACTGGATGCAAGGCTGCGCGCGCTGCGGAGAGAGTGGGGCGTGAAC	2700	
Db	2641	GGCAGCTGCTACTGGATGCAAGGCTGCGCGCGCTGCGGAGAGAGTGGGGCGTGAAC	2700	
Qy	2701	TGCTGGACCTGCTCCAGGCCCCCGGAGAGGCGCTGTGTGTGCTGTCAGGGGCTCCCCAC	2760	
Db	2701	TGCTGGACCTGCTCCAGGCCCCCGGAGAGGCGCTGTGTGTGCTGTCAGGGGCTCCCCAC	2760	
Qy	2761	CAGGTGACGAGGCGCTGGTGAGCACAGTCAAGCGTCACTCAGCACTTCTCTCCCTGAGACC	2820	
Db	2761	CAGGTGACGAGGCGCTGGTGAGCACAGTCAAGCGTCACTCAGCACTTCTCTCCCTGAGACC	2820	
Qy	2821	TCTGCCCTCTCTGTCAAGCTCTGCCAACAGGGA CCGAGCTTCCCTCTGACTGCCACTG	2880	
Db	2821	TCTGCCCTCTCTGTCAAGCTCTGCCAACAGGGA CCGAGCTTCCCTCTGACTGCCACTG	2880	
Qy	2881	CTTTATGCCAGATGGA CTGGGCTGTGTTCACAGCAGTGAAGTGGCGTGGGACATTA	2940	
Db	2881	CTTTATGCCAGATGGA CTGGGCTGTGTTCACAGCAGTGAAGTGGCGTGGGACATTA	2940	
Qy	2941	CAGGAGGCCAAATAGAGGATGCTAGGTGTCTTGGGATCGGGGTGGGACAGGTAGACAG	3000	
Db	2941	CAGGAGGCCAAATAGAGGATGCTAGGTGTCTTGGGATCGGGGTGGGACAGGTAGACAG	3000	
Qy	3001	GTGCTACGCCAGGCACAACTTTCAGCAGGGGATGGCGCTAGGGACCTTGGGGATTTCTGG	3060	
Db	3001	GTGCTACGCCAGGCACAACTTTCAGCAGGGGATGGCGCTAGGGACCTTGGGGATTTCTGG	3060	
Qy	3061	TCAACCCACAGCACCACTCTGGGCACAGCAGGGCACTGTGTTCCCTCCCTCTTAAG	3120	
Db	3061	TCAACCCACAGCACCACTCTGGGCACAGCAGGGCACTGTGTTCCCTCCCTCTTAAG	3120	
Qy	3121	CCAAACAACACAGTGCACCAAGCTCACACCTGTCTTCTCAGGCTGGCATCTCCCCAC	3180	
Db	3121	CCAAACAACACAGTGCACCAAGCTCACACCTGTCTTCTCAGGCTGGCATCTCCCCAC	3180	
Qy	3181	CTGTGCGCTTTTATGTA CAGG	3202	
Db	3181	CCTGTGCGCTTTTATGTA CAGG	3202	

## RESULT 2

RESULT 2  
 US-10-122-013-1  
 ; Sequence 1, Application US/10122013  
 ; Publication NO. US20030077614A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Christiano, Angela  
 ; TITLE OF INVENTION: NUCLEIC ACIDS FOR INHIBITING HAIRLESS PROTEIN EXPRESSION AND METH  
 ; TITLE OF INVENTION: USE THEREOF  
 ; FILE REFERENCE: 0575/62637A  
 ; CURRENT APPLICATION NUMBER: US/10/122,013  
 ; CURRENT FILING DATE: 2002-04-12  
 ; NUMBER OF SEQ ID NOS: 27  
 ; SOFTWARE: PatentIn version 3.0

QY 181 CTCCTCTGGGAGCCAGACACACTGTGCTGACCTCTGCTGGCCGCTTGTCTCCCGAGGC 240  
DB 181 CTCCTCTGGGAGCCAGACACACTGTGCTGACCTCTGCTGGCCGCTTGTCTCCCGAGGC 240  
QY 241 CTTGTTTCACTCTTTGGCAACGCTCTGGGCTGGGCCAGGCGATGGGAACCTTGGGTACACAG 300  
DB 241 CTTGTTTCACTCTTTGGCAACGCTCTGGGCTGGGCCAGGCGATGGGAACCTTGGGTACACAG 300  
QY 301 CTGGGGCCACAGCAACACAAAGGTGCCCCCTCTCTGAGCGCTGTCTACCCAGCGGGGC 360  
DB 301 CTGGGGCCACAGCAACACAAAGGTGCCCCCTCTCTGAGCGCTGTCTACCCAGCGGGGC 360  
QY 361 TGCTGTTTCACTCTACCCACCTCTAAAGGTGGGATCTTTGGCCCTTGTGGGAAGTGCCAG 420  
DB 361 TGCTGTTTCACTCTACCCACCTCTAAAGGTGGGATCTTTGGCCCTTGTGGGAAGTGCCAG 420  
QY 421 GAGGCCCTGGAGGGGGTGCAGTGGAGCCAGCGAACCCAGCGAGGAAGTGAACAAGGOC 480  
DB 421 GAGGCCCTGGAGGGGGTGCAGTGGAGCCAGCGAACCCAGCGAGGAAGTGAACAAGGOC 480  
QY 481 TCTGGCCCAAGGCGCTGTCCCCCAGCCACACACAAAGCTGAAGAAGCATGGCTCA 540  
DB 481 TCTGGCCCAAGGCGCTGTCCCCCAGCCACACACAAAGCTGAAGAAGCATGGCTCA 540  
QY 541 CGGCACTCGAGAGCAGTTTGAATGTTCCAGCGGCTGCTGAGGTGAGGAGGCGGTT 600  
DB 541 CGGCACTCGAGAGCAGTTTGAATGTTCCAGCGGCTGCTGAGGTGAGGAGGCGGTT 600  
QY 601 GCTCGGCTCCGGGCCCTCAAAAGGCGAGCAGCCCGAGGTCAGGAGCAATGGGCAGT 660  
DB 601 GCTCGGCTCCGGGCCCTCAAAAGGCGAGCAGCCCGAGGTCAGGAGCAATGGGCAGT 660  
QY 661 CCAGCCCCAAGCGCCACCGGACCTTTCCAGGCACTGCAGAACAGGGGCTGGGGGT 720  
DB 661 CCAGCCCCAAGCGCCACCGGACCTTTCCAGGCACTGCAGAACAGGGGCTGGGGGT 720  
QY 721 TGGCAGGAGTTTGGGACACATCGATGAGGAACAGGATGTGCACTCGGACAGCATGAT 780  
DB 721 TGGCAGGAGTTTGGGACACATCGATGAGGAACAGGATGTGCACTCGGACAGCATGAT 780  
QY 781 GAGCAAGAGAACCCCAAGATGGCCAGCCAGTCTCCAGGACCCGGGACTTCCAGGACATA 840  
DB 781 GAGCAAGAGAACCCCAAGATGGCCAGCCAGTCTCCAGGACCCGGGACTTCCAGGACATA 840  
QY 841 CCATGCTGCTTCTCCCTGCAAACTGCTCAATGCTCAAAAGTTGTCGCCAGCAGCTGGA 900  
DB 841 CCATGCTGCTTCTCCCTGCAAACTGCTCAATGCTCAAAAGTTGTCGCCAGCAGCTGGA 900  
QY 901 GAGGAGGAGGCGCACGCTGCTCAGCAAGTGGGAGATCGCCCTCTGGGAGGGAG 960  
DB 901 GAGGAGGAGGCGCACGCTGCTCAGCAAGTGGGAGATCGCCCTCTGGGAGGGAG 960  
QY 961 CTGACAGAGGAGGAGACACAGCCACCACTCAGCTCTGAGGAAGGCCAGGCTCGGCG 1020  
DB 961 CTGACAGAGGAGGAGACACAGCCACCACTCAGCTCTGAGGAAGGCCAGGCTCGGCG 1020  
QY 1021 CCTCAGAGCGGCTCAGCAGAGGCTGCGCAGCACCTGCTCAGTGTGTTGGGGAACGA 1080  
DB 1021 CCTCAGAGCGGCTCAGCAGAGGCTGCGCAGCACCTGCTCAGTGTGTTGGGGAACGA 1080  
QY 1081 CTGTGCGGCTGCTGCGGAGGAGCGGAGGCGCTTGGCTTGGGCCCGAGCGGGAAGGCCAA 1140  
DB 1081 CTGTGCGGCTGCTGCGGAGGAGCGGAGGCGCTTGGCTTGGGCCCGAGCGGGAAGGCCAA 1140  
QY 1141 GGGCCAGCGGTGACAGGGGACAGCCAGGCAATCCAGCTGCTGAGCGGTTGCCACCAT 1200  
DB 1141 GGGCCAGCGGTGACAGGGGACAGCCAGGCAATCCAGCTGCTGAGCGGTTGCCACCAT 1200  
QY 1201 GGACTCTTACACCCACTGGGATGTTCCCGCTGACGACCGGCTGTGTGGGCTGT 1260  
DB 1201 GGACTCTTACACCCACTGGGATGTTCCCGCTGACGACCGGCTGTGTGGGCTGT 1260

QY 1261 GGTCTGTGTGGCAGGCACTGGGCGGCGCCAGGAGAAAGCAGGCTTTTCAGGAGCAGTCCGCG 1320  
DB 1261 GGTCTGTGTGGCAGGCACTGGGCGGCGCCAGGAGAAAGCAGGCTTTTCAGGAGCAGTCCGCG 1320  
QY 1321 GAGGAGTGCACCGCAGAGGCGCGGACGCTGCTGCTTCCCTGATGCTGACCCAGTTTGTTC 1380  
DB 1321 GAGGAGTGCACCGCAGAGGCGCGGACGCTGCTGCTTCCCTGATGCTGACCCAGTTTGTTC 1380  
QY 1381 TCCAGCCAGGCTTTTGGCAGAGCTGAGCACTGCAATGCAACAGGCTTGGGTCAAGTTTGTAT 1440  
DB 1381 TCCAGCCAGGCTTTTGGCAGAGCTGAGCACTGCAATGCAACAGGCTTGGGTCAAGTTTGTAT 1440  
QY 1441 ATCCGGGGGCACTGCCCCCTGCAAGCTGATGCCGGGTATGGGCCCCCGGGGATGACGAGC 1500  
DB 1441 ATCCGGGGGCACTGCCCCCTGCAAGCTGATGCCGGGTATGGGCCCCCGGGGATGACGAGC 1500  
QY 1501 CAGCAGAAAGGATCAACACAGAAACGCCCCAACTCCACACCTTCTGCAATGGCGAC 1560  
DB 1501 CAGCAGAAAGGATCAACACAGAAACGCCCCAACTCCACACCTTCTGCAATGGCGAC 1560  
QY 1561 ACCCAAGGACCAAGAGCATCAAGAGGAGACCCCCGATTCGCTGAGACCCCGAGCAGAG 1620  
DB 1561 ACCCAAGGACCAAGAGCATCAAGAGGAGACCCCCGATTCGCTGAGACCCCGAGCAGAG 1620  
QY 1621 GACCGTGTGGCGGAGGCGCCCTGCTTGTCTCTCTGCGAACTGCTGCTTCTTACC 1680  
DB 1621 GACCGTGTGGCGGAGGCGCCCTGCTTGTCTCTCTCTGCGAACTGCTGCTTCTTACC 1680  
QY 1681 GCGGTCAAACTCTGCTTGGGCCATGAGCGAATACACATGGGCTTTCGCCCCGCTTCTGCG 1740  
DB 1681 GCGGTCAAACTCTGCTTGGGCCATGAGCGAATACACATGGGCTTTCGCCCCGCTTCTGCG 1740  
QY 1741 GCGGTCAAACTCTGCTTGGGCCATGAGCGAATACACATGGGCTTTCGCCCCGCTTCTGCG 1800  
DB 1741 GCGGTCAAACTCTGCTTGGGCCATGAGCGAATACACATGGGCTTTCGCCCCGCTTCTGCG 1800  
QY 1801 GAAAGGAGATTCAGGAGAAAGCCCTGGGCGGGGCTTCGAGCTGCGGCCCGGCTTCTGCGC 1860  
DB 1801 GAAAGGAGATTCAGGAGAAAGCCCTGGGCGGGGCTTCGAGCTGCGGCCCGGCTTCTGCGC 1860  
QY 1861 AAGGCGCTGGGCGCTGCGCCCTCTCTCCAGTGGGCGCGGCTGCGCTCCCGAGGGGCTTTCG 1920  
DB 1861 AAGGCGCTGGGCGCTGCGCCCTCTCTCCAGTGGGCGCGGCTGCGCTCCCGAGGGGCTTTCG 1920  
QY 1921 CTGTGCTGTCAGGAGCGCCAGCCTTGGCCCTGCGGCTGGCTTCCACCTCTTCCAGGAGCAG 1980  
DB 1921 CTGTGCTGTCAGGAGCGCCAGCCTTGGCCCTGCGGCTGGCTTCCACCTCTTCCAGGAGCAG 1980  
QY 1981 TGGAGGCGAGGCGCAGCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2040  
DB 1981 TGGAGGCGAGGCGCAGCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2040  
QY 2041 TGGGGAAGAGAGCTTGGGGAACCTTGGGGAACCTTGGGGAACCTTGGGGAACCTTGGGGA 2100  
DB 2041 TGGGGAAGAGAGCTTGGGGAACCTTGGGGAACCTTGGGGAACCTTGGGGAACCTTGGGGA 2100  
QY 2101 CTTCCCGAGCGCAGCCTGCGGAGCAACAACTTCTGGGAGGCTTCTCTGCGGCTGAG 2160  
DB 2101 CTTCCCGAGCGCAGCCTGCGGAGCAACAACTTCTGGGAGGCTTCTCTGCGGCTGAG 2160  
QY 2161 CTTGCGCCAAAGTCAGAGGAGGCTCTGCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2220  
DB 2161 CTTGCGCCAAAGTCAGAGGAGGCTCTGCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2220  
QY 2221 GACACAGCAGGCTGAGGAACCTTAGCTGCGAGTCCACTTCCGAGTACTCGGCCCTC 2280  
DB 2221 GACACAGCAGGCTGAGGAACCTTAGCTGCGAGTCCACTTCCGAGTACTCGGCCCTC 2280  
QY 2281 CATGGAAACTCAACCTGGCTTCTTACTCCCGGCGCTTGGCTGCTGCTGCTGCTGCTGCTGCTG 2340  
DB 2281 CATGGAAACTCAACCTGGCTTCTTACTCCCGGCGCTTGGCTGCTGCTGCTGCTGCTGCTGCTG 2340  
QY 2341 CCCCAGCTCTGGGAGGCTTATGCTGAGCGCCGACCGGGGACACCTTGGGGAACCAAGAAC 2400

[illegible]

121

[illegible]

## RESULT 5

```

RESULT 5
US-09-103-840A-1/c
; Sequence 1, Application US/09103840A
; Patent No. 6294328
; GENERAL INFORMATION:
; APPLICANT: FLEISCHMAN, Robert D.
; APPLICANT: WHITE, Owen R.
; APPLICANT: FRASER, Claire M.
; APPLICANT: VENTER, John C.
; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
; TITLE OF INVENTION: TUBERCULOSIS
; FILE REFERENCE: 24366-20007.00
; CURRENT APPLICATION NUMBER: US/09/103,840A
; CURRENT FILING DATE: 1998-06-24
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: Patent In Ver. 2.1

```

```

Query Match          1.5%;   Score 48.2;   DB 3;   Length 4411529;
Best Local Similarity 49.4%;
Matches 125; Conservative 0; Mismatches 128; Indels 0; Gaps 0;

```

[illegible][illegible]

Qy 1173 TCCAGCTGCTGCAGCGGTTGCCACCATGACTCTTCAACACCCACTGCGGATGTCCCGG 1232

Db 3948635 GTGCCGCGGTGCCCGCGGTGCCGCCGTTGAAGTTGGCCCGCGCTACCGCGGCGCCCG 3948576

[illegible]

ADDRESSEE: FOLEY & LARDNER  
STREET: 1800 Diagonal Road, Suite 500

2581 CGCGCTTTTCCAGATGTTGTCGCCGCGCGGCGAGCGCCCTGGAGCCTGGGCGCCCA 2640  
Db CGCGCTTTTCCAGATGTTGTCGCCGCGCGGCGAGCGCCCTGGAGCCTGGGCGCCCA 2640  
2641 GCGAGCTGCTACCTGGATGCGAGGCTGCGCGCGCGCCCTGGGAGGAGTGGGCGTGAGC 2700  
Db GCGAGCTGCTACCTGGATGCGAGGCTGCGCGCGCGCCCTGGGAGGAGTGGGCGTGAGC 2700  
2701 TGCTGGACCTTCCAGGCGCCCGGAGAGCGCGTGTGCTGCGAGGGGCTCCCGAC 2760  
Db TGCTGGACCTTCCAGGCGCCCGGAGAGCGCGTGTGCTGCGAGGGGCTCCCGAC 2760  
2761 CAGGTGACGGGCTGCTGAGACAGTCACTGAGCTTCTCTCCCTCCAGAC 2820  
Db CAGGTGACGGGCTGCTGAGACAGTCACTGAGCTTCTCTCCCTCCAGAC 2820  
2821 TCTGCGCTTCTGCTCAGTCTGCCACAGGAGCCAGCTTCCCTGACTGCCACCTG 2880  
Db TCTGCGCTTCTGCTCAGTCTGCCACAGGAGCCAGCTTCCCTGACTGCCACCTG 2880  
2881 CTTTATGCCAGTGGTGGCTGTGTTCCAGAGTGAAGTGGCGTGGGACATTA 2940  
Db CTTTATGCCAGTGGTGGCTGTGTTCCAGAGTGAAGTGGCGTGGGACATTA 2940  
2941 CAGGAGGCCAATAGAGGATGCTAGGTGCTGCGGATCGGGTGGGACAGTAGACAG 3000  
Db CAGGAGGCCAATAGAGGATGCTAGGTGCTGCGGATCGGGTGGGACAGTAGACAG 3000  
3001 GTGCTCAGCGCAGGACCACTTACAGAGGGATGGGCTAGGGAGTGGGATTTCTGG 3060  
Db GTGCTCAGCGCAGGACCACTTACAGAGGGATGGGCTAGGGAGTGGGATTTCTGG 3060  
3061 TCAACCCACAGACCACTTCTGGGACAGGAGGCTCTGTTCCCTCCCTTAAG 3120  
Db TCAACCCACAGACCACTTCTGGGACAGGAGGCTCTGTTCCCTCCCTTAAG 3120  
3121 CCAACACACAGTGCACCAAGCTCACACCTGTCTTCTCAGGCTGGCATCTCCCGCAC 3180  
Db CCAACACACAGTGCACCAAGCTCACACCTGTCTTCTCAGGCTGGCATCTCCCGCAC 3180  
3181 CCGTGGCCTTTATGTACAGG 3202  
Db CCGTGGCCTTTATGTACAGG 3202

RESULT 2  
US-08-447-965A-1/c  
; Sequence 1, Application US/08447965A  
; Patent No. 5776692  
; GENERAL INFORMATION:  
; APPLICANT: El-Zaatari, Fouad A. K.  
; APPLICANT: Graham, David Y.  
; APPLICANT: Naeer, Saleh  
; TITLE OF INVENTION: Mycobacterial Genus-Specific DNA Probe  
; TITLE OF INVENTION: and its Expressed Product  
; NUMBER OF SEQUENCES: 2  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Fulbright & Jaworski L.L.P.  
; STREET: 801 Pennsylvania Ave. N.W.  
; CITY: Washington, D.C.  
; COUNTRY: U.S.A.  
; ZIP: 20004  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/447,965A  
; FILING DATE: 23-MAY-1995  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Davis, Peter J.

REGISTRATION NUMBER: 36,119  
REFERENCE/DOCKET NUMBER: 311,019  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-662-4669  
TELEFAX: 202-662-4643  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1402 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
HYPOTHETICAL: NO  
US-08-447-965A-1  
Query Match 1.6%; Score 52; DB 1; Length 1402;  
Best Local Similarity 46.2%; Pred. No. 0.014;  
Matches 169; Conservative 0; Mismatches 197; Indels 0; Gaps 0;  
QY 2401 CTCTGTGGAGGTGGCGACCTGTGAGCAATCTTCTGGTGGTGCATGCGACACACCACTGCT 2460  
Db 617 CTGTGCGTGAAGCTGCCGACGCGGACACCGTCAAGCTTTCGCGCCANCCAGACCTACGCC 558  
QY 2461 GCCTGGCACCGGACACAGAAAGACTTCTTTTTCAGGCTGAGCGGAGGGGCTCTGCTCT 2520  
Db 557 GACGGCACCGTGTCAAGTGGGATCAGCCACCGCTGCCGCGGCGGAGCCCGAGCAC 498  
QY 2521 CCGGCGACCGAGTGTGAGCACTGTGTGGCACGCTTTCGCGGACAGGACGCCAGCGCATC 2580  
Db 497 CCGGCGGCGACGCTGCGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 438  
QY 2581 CCGCGCTTTTCTCCAGATGTTGTCGCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 2640  
Db 437 CCGCGCGCGCGCAACGCGCGCGTGTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCT 378  
QY 2641 GCGAGCTGTACTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2700  
Db 377 GGGCATGCGGATGCGGCTGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCT 318  
QY 2701 TGCTGACCGCTGCTCAGGCGCGCGGAGAGGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2760  
Db 317 GTGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 258  
QY 2761 CAGGTG 2766  
Db 257 CCGGTG 252

RESULT 3  
US-09-252-991A-1027  
; Sequence 1027, Application US/09252991A  
; Patent No. 6551795  
; GENERAL INFORMATION:  
; APPLICANT: Marc J. Rubenfield et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS  
; FILE REFERENCE: 107196.136  
; CURRENT APPLICATION NUMBER: US/09/252,991A  
; CURRENT FILING DATE: 1999-02-18  
; PRIOR APPLICATION NUMBER: US 60/074,788  
; PRIOR FILING DATE: 1998-02-18  
; PRIOR APPLICATION NUMBER: US 60/094,190  
; PRIOR FILING DATE: 1998-07-27  
; NUMBER OF SEQ ID NOS: 33142  
; SEQ ID NO 1027  
; LENGTH: 1257  
; TYPE: DNA  
; ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-1027  
Query Match 1.6%; Score 50.6; DB 4; Length 1257;  
Best Local Similarity 45.5%; Pred. No. 0.027;  
Matches 179; Conservative 0; Mismatches 214; Indels 0; Gaps 0;



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 29, 2004, 13:21:41 ; Search time 216 Seconds  
(without alignments)  
8226.633 Million cell updates/sec

Title: US-10-024-368-1

Perfect score: 3202

Sequence: 1 ttttactacaagatccagag.....tctgcccctttatgacagg 3202

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents NA:\*

1: /cgn2\_6/ptodata/2/ina/5A COMB.seq:\*

2: /cgn2\_6/ptodata/2/ina/5B COMB.seq:\*

3: /cgn2\_6/ptodata/2/ina/6A COMB.seq:\*

4: /cgn2\_6/ptodata/2/ina/6B COMB.seq:\*

5: /cgn2\_6/ptodata/2/ina/PTUS COMB.seq:\*

6: /cgn2\_6/ptodata/2/ina/backfiles1.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3202	100.0	3202	4	US-09-287-354-1
2	52	1.6	1402	1	US-08-447-965A-1
3	50.6	1.6	1257	4	US-09-252-991A-1027
4	49.8	1.6	1692	4	US-09-252-991A-978
5	48.2	1.5	4411529	3	US-09-103-840A-1
6	47.6	1.5	7218	1	US-08-232-463-14
7	46	1.4	498	4	US-09-252-991A-1030
8	45.6	1.4	80161	3	US-09-036-987A-1
9	45.6	1.4	80161	3	US-09-370-700-1
10	45.6	1.4	80161	4	US-09-603-207-1
11	44.8	1.4	430	4	US-09-621-976-16656
12	44.4	1.4	1512	4	US-09-252-991A-14111
13	44.4	1.4	1866	4	US-09-252-991A-14118
14	44.4	1.4	4287	4	US-09-252-991A-14160
15	43.8	1.4	30001	1	US-08-125-468-1
16	43.8	1.4	30001	2	US-08-474-933-1
17	43.6	1.4	5467	1	US-07-745-206A-12
18	43.6	1.4	5467	2	US-08-311-363-12
19	43.6	1.4	5467	4	US-09-620-312D-41
20	43.4	1.4	1254	4	US-09-252-991A-7536
21	43.4	1.4	1989	4	US-09-252-991A-7805
22	43.4	1.4	2109	4	US-09-370-838-153
23	43.4	1.4	2313	4	US-09-252-991A-7945
24	42.8	1.3	367	4	US-09-072-596-249
25	42.8	1.3	367	4	US-09-072-967-254
26	42	1.3	540	4	US-09-252-991A-10333
27	42	1.3	891	4	US-09-252-991A-10527

RESULT 1

US-09-287-354-1

; Sequence 1, Application US/09287354

; Patent No. 6348348

; GENERAL INFORMATION:

; APPLICANT: THOMPSON, Catherine C.

; TITLE OF INVENTION: HUMAN HAIRLESS GENE AND PROTEIN

; FILE REFERENCE: Thompson-20263/0243435

; CURRENT APPLICATION NUMBER: US/09/287,354

; EARLIER FILING DATE: 1999-04-07

; EARLIER APPLICATION NUMBER: US 60/080,888

; EARLIER FILING DATE: 1998-04-07

; NUMBER OF SEQ ID NOS: 10

; SOFTWARE: Patentin Ver. 2.0

; SEQ ID NO 1

; LENGTH: 3202

; TYPE: DNA

; ORGANISM: Homo sapiens

US-09-287-354-1

Query Match 100.0%; Score 3202; DB 4; Length 3202;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 3202; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	TTTTACTACAAGATCCGAGATTC	1320	4	US-09-252-991A-10535	Sequence 10535, A
		TTTTACTACAAGATCCGAGATTC	1473	4	US-09-252-991A-10639	Sequence 10639, A
Db	1	TTTTACTACAAGATCCGAGATTC	1723	4	US-09-614-034-187	Sequence 187, App
		TTTTACTACAAGATCCGAGATTC	1723	4	US-09-016-434-1182	Sequence 1182, Ap
QY	61	CTTGGGTTGTTTGGCTTAAACTCTGTGTGGCACTCTGCAGAGAGCGGGGAGGCGGAACGC	120			
Db	61	CTTGGGTTGTTTGGCTTAAACTCTGTGTGGCACTCTGCAGAGAGCGGGGAGGCGGAACGC	120			
QY	121	CCTTCACTGCACAGAGGATGGAGAGATGGGAGCTGGCCGGCAGACAGATCCTTGGCCGC	180			
Db	121	CCTTCACTGCACAGAGGATGGAGAGATGGGAGCTGGCCGGCAGACAGATCCTTGGCCGC	180			
QY	181	CTCTTCTGGGGCAGCCAGACACTGTGTCCCTTGAACCTCTCTGGCCCGCTTGTCCCCCAGGC	240			
Db	181	CTCTTCTGGGGCAGCCAGACACTGTGTCCCTTGAACCTCTCTGGCCCGCTTGTCCCCCAGGC	240			
QY	241	CTTGTTCATACTCTTGGCAACGTCCTGGCTGGGGCAGGCGCATGGGAACCTTGGGTACCA	300			
Db	241	CTTGTTCATACTCTTGGCAACGTCCTGGCTGGGGCAGGCGCATGGGAACCTTGGGTACCA	300			
QY	301	CTGGGGCCACAGCAACACCAAGTGCCCTCTCTGAGCGCGCTGTACCCACCGGGGC	360			
Db	301	CTGGGGCCACAGCAACACCAAGTGCCCTCTCTGAGCGCGCTGTACCCACCGGGGC	360			
QY	361	TGCTGTTCATCTACCCACCCACTAAAGTTGGGATCTTGGCCCTTGTGGGAAGTGCAC	420			

ALIGNMENTS





GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model  
Run on: June 29, 2004, 13:08:50 ; Search time 7850 seconds  
(without alignments)  
12180.726 Million cell updates/sec

Title: US-10-024-368-1  
Perfect score: 3202  
Sequence: 1 ttttactacaagatccgag.....tgtgccttttatgacagg 3202

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

EST:\*

- 1: em\_estba:\*
- 2: em\_esthum:\*
- 3: em\_estin:\*
- 4: em\_estm:\*
- 5: em\_estov:\*
- 6: em\_estpl:\*
- 7: em\_estro:\*
- 8: em\_htc:\*
- 9: gb\_est1:\*
- 10: gb\_est2:\*
- 11: gb\_htc:\*
- 12: gb\_est3:\*
- 13: gb\_est4:\*
- 14: gb\_est5:\*
- 15: em\_estfun:\*
- 16: em\_estcom:\*
- 17: em\_gss\_hum:\*
- 18: em\_gss\_inv:\*
- 19: em\_gss\_pln:\*
- 20: em\_gss\_vrt:\*
- 21: em\_gss\_fun:\*
- 22: em\_gss\_mam:\*
- 23: em\_gss\_mus:\*
- 24: em\_gss\_pro:\*
- 25: em\_gss\_rod:\*
- 26: em\_gss\_phg:\*
- 27: em\_gss\_vrl:\*
- 28: gb\_gss1:\*
- 29: gb\_gss2:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1273	39.8	2969	11	AK076280 Mus muscu
2	781.4	24.4	1094	12	BM799906 AGENCOURT
3	764.8	23.9	933	13	BQ961259 AGENCOURT
4	736.2	23.0	811	12	BG739858 602630590

5	686.2	21.4	1062	12	BI197548
6	665.8	20.8	837	13	BX341278
7	659.4	20.6	869	12	BG743219
8	657	20.5	989	10	BF316119
9	636.2	19.9	780	12	BI827477
10	628.8	19.6	725	12	BG677197
11	621.2	19.4	858	12	BG330970
12	609	19.0	978	10	BF346259
13	594.4	18.6	951	12	BI261186
14	574	17.9	586	14	CB154516
15	566	17.7	802	10	BE276310
16	556.2	17.4	1068	12	BG678990
17	554	17.3	861	10	BE784838
18	549.6	17.2	618	10	BE410741
19	523.2	16.3	867	13	BQ951435
20	520	16.2	775	12	BG923878
21	481.2	15.0	855	13	BX415340
22	440	13.7	654	12	BG963522
23	439.6	13.7	921	13	BQ925145
24	428.4	13.4	783	13	BQ444010
25	424.4	13.3	554	10	BF077353
26	424.2	13.2	732	12	BI655486
27	418.6	13.1	875	12	BI687595
28	418.6	13.1	909	10	BF528824
29	394	12.3	700	13	BQ443965
30	367.2	11.5	623	10	BB612436
31	360.4	11.3	515	14	CD548907
32	353.4	11.0	789	13	BQ217662
33	349.8	10.9	894	13	BQ217662
34	336.8	10.5	624	10	BE573367
35	330.8	10.3	471	9	AA760207
36	321	10.0	481	13	BQ307325
37	313.2	9.8	490	14	R67180
38	308.8	9.6	771	12	BI665435
39	307.6	9.6	543	14	CB612355
40	306	9.6	916	13	BUI96087
41	301.6	9.4	867	12	BI655258
42	285	8.9	389	10	BF225274
43	283.4	8.9	820	28	AQ781018
44	268.8	8.4	474	12	BG992856
45	266.8	8.3	311	10	AW354018

ALIGNMENTS

RESULT 1	AK076280	2969 bp	linear	HTC 18-SEP-2003
LOCUS	Mus musculus 0 day neonate skin cDNA, RIKEN full-length enriched library, clone:4632409G21 product:hairless, full insert sequence.			
DEFINITION	AK076280	GI:26096706		
ACCESSION	AK076280.1	GI:26096706		
VERSION	AK076280.1	GI:26096706		
KEYWORDS	HTC; CAP trapper.			
SOURCE	Mus musculus (house mouse)			
ORGANISM	Mus musculus			
REFERENCE	1			
AUTHORS	Carninci, P. and Hayashizaki, Y.			
TITLE	High-efficiency full-length cDNA cloning			
JOURNAL	Meth. Enzymol. 303, 19-44 (1999)			
MEDLINE	99279253			
PUBMED	10349636			
REFERENCE	2			
AUTHORS	Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Komoto, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.			
TITLE	Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes			
JOURNAL	Genome Res. 10 (10), 1617-1630 (2000)			
MEDLINE	20499374			
PUBMED	11042159			
REFERENCE	3			



Db 2836 GACACACAGGCTGGAGAACTAGCTGCAGTCTGCCACTTCCGGAGTACTGCGCCCTC 2895  
Qy 2281 CATGMAAACTCAACCTGGCTTCTTACCTCCACCGGGCCCTTCCCTGCTGCTGAG 2340  
Db 2896 CATGMAAACTCAACCTGGCTTCTTACCTCCACCGGGCCCTTCCCTGCTGCTGAG 2955  
Qy 2341 CCCAGCTCTGGGAGCTATGCTGTGAGCCCGGACCGGGGACACTGGGGACCAAGAAC 2400  
Db 2956 CCCAGCTCTGGGAGCTATGCTGTGAGCCCGGACCGGGGACACTGGGGACCAAGAAC 3015  
Qy 2401 CTCTGTGTGGAGTGGCGGACCTGTGTGAGCACTTCTGTGTGATGCTGCTGCTGCT 2460  
Db 3016 CTCTGTGTGGAGTGGCGGACCTGTGTGAGCACTTCTGTGTGATGCTGCTGCTGCT 3075  
Qy 2461 GCCTGGACCGGGACAGAAAGACTTCTTTCAGGCTTGGAGGGGAGGGCTTGGTCT 2520  
Db 3076 GCCTGGACCGGGACAGAAAGACTTCTTTCAGGCTTGGAGGGGAGGGCTTGGTCT 3135  
Qy 2521 CGGGCAGCAGCTGAGCACTGTGTGGACGCTGTTCGGGGCAGAGACCGCCAGCGCATC 2580  
Db 3136 CGGGCAGCAGCTGAGCACTGTGTGGACGCTGTTCGGGGCAGAGACCGCCAGCGCATC 3195  
Qy 2581 CGCCGCTTCTCCAGATGTGTGCCCGCGGGGACAGCGCTGTGAGCTTGGCGCCCA 2640  
Db 3196 CGCCGCTTCTCCAGATGTGTGCCCGCGGGGACAGCGCTGTGAGCTTGGCGCCCA 3255  
Qy 2641 GSCAGCTGCTACTGATCAGGCTGTGGCGGCTGCGCGGCTGCGGAGGAGTGGCGCTGAGC 2700  
Db 3256 GSCAGCTGCTACTGATCAGGCTGTGGCGGCTGCGCGGCTGCGGAGGAGTGGCGCTGAGC 3315  
Qy 2701 TGTGACCTCTGCTCCAGGCCCCCGGAGAGGCGCTGCTGTGCTGCTGAGGGGCTCCCCAC 2760  
Db 3316 TGTGACCTCTGCTCCAGGCCCCCGGAGAGGCGCTGCTGTGCTGAGGGGCTCCCCAC 3375  
Qy 2761 CAGGTGACGGGCTGTGAGCAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 2820  
Db 3376 CAGGTGACGGGCTGTGAGCAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 3435  
Qy 2821 TGTGCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2880  
Db 3436 TGTGCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3495  
Qy 2881 CTTTATGCCAGATGAGTGGGCTGTGTTCCAAAGCAGTGAAGTGGCGCTGGGGACATTA 2940  
Db 3496 CTTTATGCCAGATGAGTGGGCTGTGTTCCAAAGCAGTGAAGTGGCGCTGGGGACATTA 3555  
Qy 2941 CAGGAGGCCAAA 2952  
Db 3556 CAGGAGGCCAAA 3567

RESULT 6  
ACA57497  
ID ACA57497 standard; cDNA; 538 BP.  
AC ACA57497;  
XX

10-JUN-2003 (first entry)

Human adipocyte Selected Interacting domain, SID, cDNA #584.

Human; ss; gene; prey; adipocyte; SID; selected interacting domain;  
anorectic; antidiabetic; protein-protein interaction; diabetes;  
yeast 2-hybrid assay; metabolic disorder; obesity.

Homo sapiens.

WO200286122-A2.

31-OCT-2002.

14-MAR-2002; 2002WO-EP003768.

XX PR 14-MAR-2001; 2001US-0275734P.  
XX (HYBR-) HYBRIGENICS.  
XX PA Legrain P, Daviet L;  
XX PI WPI; 2003-103412/09.  
XX DR P-PSDB; ABU70953.  
XX PT New complex between two interacting proteins in adipocyte cells, useful  
PT for identifying selected interacting domains that modulate protein  
PT interactions, or for preventing or treating metabolic disorders such as  
PT obesity or diabetes.  
XX PS Claim 7; Page 297; 382pp; English.  
XX CC The invention relates to a complex between two interacting proteins in  
CC adipocyte cells, given in the specification. The proteins are identified  
CC by selecting a bait protein from a known adipocyte marker and then  
CC performing a yeast 2-hybrid selection to isolate prey proteins encoded by  
CC members of an adipocyte cDNA library. The proteins are designated SID  
CC (RTM) (selected interacting domains) proteins. Also included are a  
CC polynucleotide encoding a polypeptide in the adipocyte cells, a  
CC recombinant host cell expressing at least one of the interacting  
CC polypeptides of the complex, selecting a modulating compound in adipocyte  
CC cells, a SID (RTM) polypeptide comprising any of the 738 amino acid  
CC sequences given in the specification (including its fragment or variant),  
CC a SID (RTM) polynucleotide comprising any of the 738 nucleotide sequences  
CC given in the specification (including its fragment or variant), a vector  
CC comprising the SID (RTM) polynucleotide, a recombinant host cell  
CC comprising the vector, a protein chip comprising the polypeptides and a  
CC record comprising all or part of the data, listed in the specification.  
CC The complex, polypeptides, polynucleotides and compounds are useful for  
CC preventing or treating metabolic disorders such as obesity or diabetes.  
CC The polynucleotides are useful as probes or primers. The complex is  
CC particularly useful for identifying selected interacting domains (SID  
CC (RTM)) for screening drugs that modulate the protein interaction, thus  
CC exhibiting the therapeutic effect. The present sequence encodes a SID  
CC (prey) protein of the invention  
XX SQ Sequence 538 BP; 105 A; 170 C; 188 G; 75 T; 0 U; 0 Other;

Query Match 16.7%; Score 533.2; DB 7; Length 538;  
Best Local Similarity 99.4%; Pred. No. 1.3e-98;  
Matches 535; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 804 CCAGGCGAGTCTCCAGGACCCGGGACTTCCAGGACATACCATGCTGCTTCTCCCTGCAAA 863  
Db 1 CCAGGCGAGTCTCCAGGACCCGGGACTTCCAGGACATACCATGCTGCTTCTCCCTGCAAA 60  
Qy 864 ACTGGCTCAATGCAAGTGTGTCCAGGACGCTGGAGAGGGAGGAGGCAAGCTGCTGCA 923  
Db 61 ACTGGCTCAATGCAAGTGTGTCCAGGACGCTGGAGAGGGAGGAGGCAAGCTGCTGCA 120  
Qy 924 CTCTAGCAAGTGGGAGATCGCTCTGGAGGGGAGCTGCAGCAGGAGGAGGACAGC 983  
Db 121 CTCTAGCAAGTGGGAGATCGCTCTGGAGGGGAGCTGCAGCAGGAGGAGGACAGC 180  
Qy 984 CACCACTCCAGCTCTGAGGAAGGCCCGAGGTCGGGCCCTGACAGCCGGCTCAGCAGG 1043  
Db 181 CACCACTCCAGCTCTGAGGAAGGCCCGAGGTCGGGCCCTGACAGCCGGCTCAGCAGG 240  
Qy 1044 CTTGCGCAAGCACCCTGCTCAGTGTGTTGGGGAGCCGACTGTGCGCTGTGCGGAGGGA 1103  
Db 241 CTTGCGCAAGCACCCTGCTCAGTGTGTTGGGGAGCCGACTGTGCGCTGTGCGGAGGGA 300  
Qy 1104 GCGGAGGCGCTGCTGGGCGCCAGCGGAGGCAAGGCCGAGCCGCTGACAGGGGACAG 1163  
Db 301 GCGGAGGCGCTGCTGGGCGCCAGCGGAGGCAAGGCCGAGCCGCTGACAGGGGACAG 360  
Qy 1164 CCCAGGCAATTCAGCTGTGTCAGCCGTTGCGCAGCTGCTTCAACACCCACTGCGC 1223

QY 61 CTTGGGTTGTTGGCTTAAACTCTGGTGGCACTTGCAGAGAGCCGGAGGCCGAAAGC 120  
Db |||||  
QY 676 CTTGGGTTGTTGGCTTAACTCTGGTGGCACTTGCAGAGAGCCGGAGGCCGAAAGC 735  
Db |||||  
QY 121 CTTCACTGCACAGAGGATGGAGATGGAGTGGCCGCGCAGCAGAACTCTTGGCCCG 180  
Db |||||  
QY 736 CTTCACTGCACAGAGGATGGAGATGGAGTGGCCGCGCAGCAGAACTCTTGGCCCG 795  
Db |||||  
QY 181 CTTCTCTGGGGAGCCAGACACTGTGGCTTGGAACTCTCTGTGCCCGCTTGTCCCCAGGC 240  
Db |||||  
QY 796 CTTCTCTGGGGAGCCAGACACTGTGGCTTGGAACTCTCTGTGCCCGCTTGTCCCCAGGC 855  
Db |||||  
QY 241 CTTGTTCATCTCTTGGCAACGCTCTGGCTGGCCAGGCGATGGAACTTGGGTACAG 300  
Db |||||  
QY 856 CTTGTTCATCTCTTGGCAACGCTCTGGCTGGCCAGGCGATGGAACTTGGGTACAG 915  
Db |||||  
QY 301 CTTGGGGCCACAGCAACCAAGGTGCCCTCTCTGTAGCCGCCCTGTCAACCAGCGGGC 360  
Db |||||  
QY 916 CTTGGGGCCACAGCAACCAAGGTGCCCTCTCTGTAGCCGCCCTGTCAACCAGCGGGC 975  
Db |||||  
QY 361 TGCTGTTCATCTCTACCAACCACTAAAGGTGGGATCTTGGCCCTTGTGGAAAGTGCCAG 1035  
Db |||||  
QY 421 GAGGCTCTGGAGGGGGTGCCAGTGGAGCCAGCAACCCAGCGAGAGTGNACAGGCC 480  
Db |||||  
QY 1036 GAGGCTCTGGAGGGGGTGCCAGTGGAGCCAGCAACCCAGCGAGAGTGNACAGGCC 1095  
Db |||||  
QY 481 TCTGGCCCTCAGGGCTGTCCCCCAGCCACACCAAGCTGAAGAACATGGCTACA 540  
Db |||||  
QY 1096 TCTGGCCCTCAGGGCTGTCCCCCAGCCACACCAAGCTGAAGAACATGGCTACA 1155  
Db |||||  
QY 541 CGGCACTCGAGAGCTTGAATGTCCAGCGGCTGCCCTGTAGGTCGAGGAGGCCGGTT 600  
Db |||||  
QY 1156 CGGCACTCGAGAGCTTGAATGTCCAGCGGCTGCCCTGTAGGTCGAGGAGGCCGGTT 1215  
Db |||||  
QY 601 GCTGGCTCTCGGGCCCTCAAAGGAGCAGCAGCCCGAGGTCAGGGAGCAATGGCAGT 660  
Db |||||  
QY 1216 GCTGGCTCTCGGGCCCTCAAAGGAGCAGCAGCCCGAGGTCAGGGAGCAATGGCAGT 1275  
Db |||||  
QY 661 CCAGCCCTCAGAGCCCAACCGAACCTTTCCAGGCACTGCAGAACAGGGGCTGGGGGT 720  
Db |||||  
QY 1276 CCAGCCCTCAGAGCCCAACCGAACCTTTCCAGGCACTGCAGAACAGGGGCTGGGGGT 1335  
Db |||||  
QY 721 TGGCAGGAGGTCGGGACACATCGATAGGAAACAGATGTGGACTCGGACAGCATGAT 780  
Db |||||  
QY 1336 TTGAGGAGGTGCGGGACACATCGATAGGAAACAGGATGTGGACTCGGGACAGCATGAT 1395  
Db |||||  
QY 781 GAGCAGAAAGGACCCCAAGATGGCCAGGCCAGTCTCCAGGACCCGGGACTTCAGGACATA 840  
Db |||||  
QY 1396 GAGCAGAAAGGACCCCAAGATGGCCAGGCCAGTCTCCAGGACCCGGGACTTCAGGACATA 1455  
Db |||||  
QY 841 CCATGCTGCTTCTCCCTGCAAAACTGGCTCAATGCAAGTTGTGCCAGGAGCTGGA 900  
Db |||||  
QY 1456 CCATGCTGCTTCTCCCTGCAAAACTGGCTCAATGCAAGTTGTGCCAGGAGCTGGA 1515  
Db |||||  
QY 901 GAGGAGGAGGACGCTGCACTCTCAGCAAGTGGGAGATGCCCTCTGGGAGGGAG 960  
Db |||||  
QY 1516 GAGGAGGAGGACGCTGCACTCTCAGCAAGTGGGAGATGCCCTCTGGGAGGGAG 1575  
Db |||||  
QY 961 CTGAGCAGGAGGAGACACAGCCCAACTCCAGCTCTGAGGAGGCCAGGCTCCGGC 1020  
Db |||||  
QY 1576 CTGAGCAGGAGGAGACACAGCCCAACTCCAGCTCTGAGGAGGCCAGGCTCCGGC 1635  
Db |||||  
QY 1021 CTTGAGCAGGCTGAGCAGAGCTGCCAAGCACTGTCTAGTGTGTTGGGGAGCGA 1080  
Db |||||  
QY 1636 CTTGAGCAGGCTGAGCAGAGCTGCCAAGCACTGTCTAGTGTGTTGGGGAGCGA 1695  
Db |||||  
QY 1081 CTTGCGCCCTGCTCGGAGGAGGAGGCCCTTGGCTGGGCCCAAGGAGGAG 1140  
Db |||||  
QY 1696 CTTGCGCCCTGCTCGGAGGAGGAGGCCCTTGGCTGGGCCCAAGGAGGAG 1755  
Db |||||

QY 1141 GGGCCAGCCGTGACAGGGGACAGCCAGGCATTCACGCTGCTGAGCCGTTGCCACCAT 1200  
Db |||||  
QY 1756 GGGCCAGCCGTGACAGAGGACAGCCAGGCATTCACGCTGCTGAGCCGTTGCCACCAT 1815  
Db |||||  
QY 1201 GGACTCTTCAACACCCACTGGCGATGTCCTGCTGAGCCACCGGCTGTGTGGCCCTGT 1260  
Db |||||  
QY 1816 GGACTCTTCAACACCCACTGGCGATGTCCTGCTGAGCCACCGGCTGTGTGTGGCCCTGT 1875  
Db |||||  
QY 1261 GGTCTGTGGCAGGCACTGGCGGGCCAGGAGAAAGCAGGCTTTCAGAGCAGTCCGGC 1320  
Db |||||  
QY 1876 GGTCTGTGGCAGGCACTGGCGGGCCAGGAGAAAGCAGGCTTTCAGAGCAGTCCGGC 1935  
Db |||||  
QY 1321 GAGAGTGCACGAGGAGGCGCGGCACTGCTGCTGCTTCCCTGATGCTGACCCAGTTTGT 1380  
Db |||||  
QY 1936 GAGAGTGCACGAGGAGGCGCGGCACTGCTGCTGCTTCCCTGATGCTGACCCAGTTTGT 1995  
Db |||||  
QY 1381 TCCAGCAGGCTTTCAGGAGGCTGAGCACTGCAATGACACAGGCTCTGGTCAAGTTTGT 1440  
Db |||||  
QY 1996 TCCAGCAGGCTTTCAGGAGGCTGAGCACTGCAATGACACAGGCTCTGGTCAAGTTTGT 2055  
Db |||||  
QY 1441 ATCCGGGGCACTGCCCCCTGCAAGCTGATGCCCGGGTATGGGCCCGCGGATGACAGC 1500  
Db |||||  
QY 2056 ATCCGGGGCACTGCCCCCTGCAAGCTGATGCCCGGGTATGGGCCCGCGGATGACAGC 2115  
Db |||||  
QY 1501 CAGCAGAAAGGATCAACACAGAAACGCCCCCAACTCCACAACTTCTGCAATGGCGAC 1560  
Db |||||  
QY 2116 CAGCAGAAAGGATCAACACAGAAACGCCCCCAACTCCACAACTTCTGCAATGGCGAC 2175  
Db |||||  
QY 1561 ACCCAGCAGGACCAAGAGCATCAAGAGGAGACCCCGATTCGCTGAGACCCAGCAGAG 1620  
Db |||||  
QY 2176 ACCCAGCAGGACCAAGAGCATCAAGAGGAGACCCCGATTCGCTGAGACCCAGCAGAG 2235  
Db |||||  
QY 1621 GACCTGTGTCGCGAGGCGCCCTGCTTGTCTTCTCTGCGAACTGCTGCTTCTACC 1680  
Db |||||  
QY 2236 GACCTGTGTCGCGAGGCGCCCTGCTTGTCTTCTCTGCGAACTGCTGCTTCTACC 2295  
Db |||||  
QY 1681 GCGGTCAAACTCTGCTTGGGCGCATGAGGAATACATATGGCTTTCGCCCTTCACTCCG 1740  
Db |||||  
QY 2296 GCGGTCAAACTCTGCTTGGGCGCATGAGCGGATACATATGGCTTTCGCCCTTCACTCCG 2355  
Db |||||  
QY 1741 GCGCTGCGCAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1800  
Db |||||  
QY 2356 GCGCTGCGCAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2415  
Db |||||  
QY 1801 GAAACGAGAGTCCAGGAGAAAGCCCTGGGGCGGGGCTTCGAGCTGGGCCCGGGCTTGGCC 1860  
Db |||||  
QY 2416 GAAACGAGAGTCCAGGAGAAAGCCCTGGGGCGGGGCTTCGAGCTGGGCCCGGGCTTGGCC 2475  
Db |||||  
QY 1861 AAGGGCTTGGGCTTGGCCCTTCTCAGTGGGCCCGGGCTTGGCTTGGCTTGGCTTGGCTTGG 1920  
Db |||||  
QY 2476 AAGGGCTTGGGCTTGGCCCTTCTCAGTGGGCCCGGGCTTGGCTTGGCTTGGCTTGGCTTGG 2535  
Db |||||  
QY 1921 CTGTGGCTGAGGAGGCGCCAGCCTTGGCCCTGCGCTTCCACCTTCCACCTTCCAGGAGCAG 1980  
Db |||||  
QY 2536 CTGTGGCTGAGGAGGCGCCAGCCTTGGCCCTGCGCTTCCACCTTCCACCTTCCAGGAGCAG 2595  
Db |||||  
QY 1981 TGGAGGAGGAGGCGCCAGCCTGCTGTGGTTCAGGGATCCAAAGGACATTCAGAGGCAACTG 2040  
Db |||||  
QY 2596 TGGAGGAGGAGGCGCCAGCCTGCTGTGGTTCAGGGATCCAAAGGACATTCAGAGGCAACTG 2655  
Db |||||  
QY 2041 TGGGGGACAGAACTCTTGGGGCACTTGGAGGCGAGGTGAGGGGCTGAGCCCTCCGGA 2100  
Db |||||  
QY 2656 TGGGGGACAGAACTCTTGGGGCACTTGGAGGCGAGGTGAGGGGCTGAGCCCTCCGGA 2715  
Db |||||  
QY 2101 CTTCCCGCAGCCAGCAGCTTGGGCGAGCACAACATTCCTGGGAGGGCTTCTCTGCGCTTGG 2160  
Db |||||  
QY 2716 CTTCCCGCAGCCAGCAGCTTGGGCGAGCACAACATTCCTGGGAGGGCTTCTCTGCGCTTGG 2775  
Db |||||  
QY 2161 CTTGCGCCAAAGTCAAGCAGGGCTCTGTCTCTGCTGACCGAGCTTTCGGGGGATGAG 2220  
Db |||||  
QY 2776 CTTGCGCCAAAGTCAAGCAGGGCTCTGTCTCTGCTGACCGAGCTTTCGGGGGATGAG 2835  
Db |||||  
QY 2221 GACACAGCAGGGTGGAGAACCTAGCTGCCAGTCTGCCACTTCGCGAGTACTGCGCCCTC 2280  
Db |||||



CC more lenient consensus cleavage site requirements and are less likely to  
CC degrade, in vivo, than hammerhead ribozymes. The catalytic nucleic acids  
CC are useful in pharmaceutical compositions for inhibiting hair production  
CC by a hair-producing cell, for inhibiting hair growth and for inhibiting  
CC the transition of a hair follicle from the anagen phase to the catagen  
CC phase. A non-human transgenic mammal is useful as a model for testing  
CC hair removal products which function by inhibiting hairless protein  
CC expression. The sequence presented is the human hairless protein cDNA  
XX  
SQ Sequence 5709 BP; 1015 A; 1950 C; 1826 G; 918 T; 0 U; 0 Other;

Query Match 99.5%; Score 3185; DB 7; Length 5709;  
Best Local Similarity 99.8%; Pred. No. 0;  
Matches 3188; Conservative 0; Mismatches 5; Indels 0; Gaps 0;  
QY 1 TTTTACTACAGGATCCAGGATTTCCAGGTTGGCAAGAGAGCCCTTGGAGCTCGGAA 60  
DB 2100 TTTTACTACAGGATCCAGGATTTCCAGGTTGGCAAGAGAGCCCTTGGAGCTCGGAA 2159  
QY 61 CCTGGTTGTTGGCTTAACTCTGCTGGCACTGACAGAGCGGGAGGCGCAAGCG 120  
DB 2160 CCTGGTTGTTGGCTTAACTCTGCTGGCACTGACAGAGCGGGAGGCGCAAGCG 2219  
QY 121 CCTTCACTGCACAGAGGATGGAGATGGAGCTGGCCGGCAGCAATCTTTGCCCG 180  
DB 2220 CCTTCACTGCACAGAGGATGGAGATGGAGCTGGCCGGCAGCAATCTTTGCCCG 2279  
QY 181 CTCCTCCCTGGGCGCAGCAGACTGTGCTGCTGGAGCTCTGCTGCTGCTGCTGCTG 240  
DB 2280 CTCCTCCCTGGGCGCAGCAGACTGTGCTGCTGGAGCTCTGCTGCTGCTGCTGCTG 2339  
QY 241 CTGTGTCATCTCTGGCAACGCTCTGGCTGGGCGCAGGCGATGGGAACTTTGGGTACCAG 300  
DB 2340 CTGTGTCATCTCTGGCAACGCTCTGGCTGGGCGCAGGCGATGGGAACTTTGGGTACCAG 2399  
QY 301 CTGGGCGCACAGCAACACCAAGGTGCCCCCTCTCTGAGCCGCTGTCCACAGCGGGCG 360  
DB 2400 CTGGGCGCACAGCAACACCAAGGTGCCCCCTCTCTGAGCCGCTGTCCACAGCGGGCG 2459  
QY 361 TGCTGTTTCATCTACCCACCACTAAAGTGGGATCTTGCCCTTGGGAAAGTGCCAG 420  
DB 2460 TGCTGTTTCATCTACCCACCACTAAAGTGGGATCTTGCCCTTGGGAAAGTGCCAG 2519  
QY 421 GAGGCTCTGAGGGGGTGCAGTGGAGCCAGCGAACCAGCGAGGAAGTGAACAGGCC 480  
DB 2520 GAGGCTCTGAGGGGGTGCAGTGGAGCCAGCGAACCAGCGAGGAAGTGAACAGGCC 2579  
QY 481 TCTGGCCCCAGGCGCTGCCCCCAGCCACCAACCAAGCTGAAGAGACATGGCTCACA 540  
DB 2580 TCTGGCCCCAGGCGCTGCCCCCAGCCACCAACCAAGCTGAAGAGACATGGCTCACA 2639  
QY 541 CGGCACCTGGAGAGTGTGAATGTCACGGGCTGCTGAGGTTCGAGGAGCGCGTT 600  
DB 2640 CGGCACCTGGAGAGTGTGAATGTCACGGGCTGCTGAGGTTCGAGGAGCGCGTT 2699  
QY 601 GCTCGGCTCCGGGCGCTCAAAAGGCGCAGCAGCCGAGGTTCAGGGAGCAATGGCGAGT 660  
DB 2700 GCTCGGCTCCGGGCGCTCAAAAGGCGCAGCAGCCGAGGTTCAGGGAGCAATGGCGAGT 2759  
QY 661 CCAGCCCCAAGCGGCCACCGGACCTTTCCAGGCACTGACAGAACCGGGGCTGGGGGT 720  
DB 2760 CCAGCCCCAAGCGGCCACCGGACCTTTCCAGGCACTGACAGAACCGGGGCTGGGGGT 2819  
QY 721 TGGCAGGAGTTCGGGACACATCATAGGAAACAGAGTGTGGATCTGGGACAGCATGAT 780  
DB 2820 TGGCAGGAGTTCGGGACACATCATAGGAAACAGAGTGTGGATCTGGGACAGCATGAT 2879  
QY 781 GAGCAGAAAGACCCCAAGATGGCCAGGCGAGTCTCCAGGACCCGGGACTTCAGACATA 840  
DB 2880 GAGCAGAAAGACCCCAAGATGGCCAGGCGAGTCTCCAGGACCCGGGACTTCAGACATA 2939  
QY 841 CCATGCTGCTTCTCCCTGCAAAACTGGCTCAATGCTCAAGTGTGCCCCAGGCGAGCTGA 900

DB 2940 CCATGCTGCTGCTTCTCCCTGCAAAACTGCTCAATGCCAAAGTTGTGCCAGGACGTGA 2999  
QY 901 GAGGAGGAGGCGCAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 960  
DB 3000 GAGGAGGAGGCGCAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3059  
QY 961 CTGACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1020  
DB 3060 CTGACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3119  
QY 1021 CCTGACGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1080  
DB 3120 CCTGACGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3179  
QY 1081 CTGCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1140  
DB 3180 CTGCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3239  
QY 1141 GGGCCAGCGCTGACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1200  
DB 3240 GGGCCAGCGCTGACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3299  
QY 1201 GGACTCTTCAACACCCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1260  
DB 3300 GGACTCTTCAACACCCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3359  
QY 1261 GGTCTGTGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1320  
DB 3360 GGTCTGTGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3419  
QY 1321 GAGGAGTGCAGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1380  
DB 3420 GAGGAGTGCAGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3479  
QY 1381 TCCAGCCAGGCTTTGGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1440  
DB 3480 TCCAGCCAGGCTTTGGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3539  
QY 1441 ATCCGGGGGCACTGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1500  
DB 3540 ATCCGGGGGCACTGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3599  
QY 1501 CAGCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1560  
DB 3600 CAGCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3659  
QY 1561 ACCCAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1620  
DB 3660 ACCCAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3719  
QY 1621 GACCGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1680  
DB 3720 GACCGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3779  
QY 1681 GCGGTCAAACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1740  
DB 3780 GCGGTCAAACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3839  
QY 1741 GCGGTCAAACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1800  
DB 3840 GCGGTCAAACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3899  
QY 1801 GAACGGAAGATTCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1860  
DB 3900 GAACGGAAGATTCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3959  
QY 1861 AAGGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCT 1920  
DB 3960 AAGGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTTTCG 4019  
QY 1921 CTGTGGCTGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1980  
DB 4020 CTGTGGCTGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 4079



Db	3061	TCAACCCCAACAGCACACTCTGGGCACAAGCAGGCACTCTGTTCCTCCCTCCCTTTAAG	312
Qy	3121	CCAAACAACACACAGTGGCCACCAAGCTCACACCTGTCTCTTCTCAGGCTGGCATCTCCCCAC	3180
Db	3121	CCAAACAACACACAGTGGCCACCAAGCTCACACCTGTCTCTTCTCAGGCTGGCATCTCCCCAC	3180
Qy	3181	CTGTGCGCCCTTTATGTACAGG	3202
Db	3181	CTGTGCGCCCTTTATGTACAGG	3202
RESULT 4			
ABX13987	ID	ABX13987 standard; cDNA; 5709 BP.	
XX	AC	ABX13987;	
XX	DT	25-FEB-2003 (first entry)	
XX	DE	Human hairless protein cDNA.	
XX	KW	Human; gene; ss; catalytic DNA; catalytic RNA; hairless protein;	
KW	KW	hair loss; atrichia; hair growth; hirsutism; catalytic nucleic acid;	
KW	KW	ribozyme; DNazymes; self-catalytic; hammerhead ribozyme; deoxy-ribozyme;	
KW	KW	catalytic core; cleavage site; pharmaceutical; hair production;	
KW	KW	hair follicle; anagen phase; catagen phase; hair removal product;	
XX	KW	depilatory.	
OS	OS	Homo sapiens.	
XX	Key	Location/Qualifiers	
PH	PH	1485..5054	
FT	FT	/*tag= a	
FT	FT	/product= "Hairless protein"	
FT	FT	/transl_except= (pos:2820..2822,aa:Leu)	
FT	FT	/transl_except= (pos:3234..3236,aa:Ser)	
XX	XX	WO200283891-A2.	
XX	XX	24-OCT-2002.	
XX	XX	12-APR-2002; 2002WO-US011683.	
XX	XX	13-APR-2001; 2001US-0283618P.	
XX	XX	(UYCO ) UNIV COLUMBIA NEW YORK.	
XX	XX	Christiano AM;	
XX	XX	WPI; 2003-093020/08.	
DR	DR	P-PSDB; ABG72775.	
XX	XX	New catalytic nucleic acid molecule that specifically cleaves Hairless	
PT	PT	Protein mRNA, useful for inhibiting hair production by a hair-producing	
PT	PT	cell, or for inhibiting transition of a hair follicle from anagen phase	
PT	PT	to catagen phase.	
XX	XX	Claim 12; Fig 1; 65pp; English.	
XX	XX	The invention discloses a new catalytic DNA or RNA molecule that	
CC	CC	specifically cleaves, or inhibits expression of, Hairless Protein mRNA	
CC	CC	which comprises a catalytic domain that cleaves mRNA at a defined	
CC	CC	consensus sequence and binding domains contiguous with the 5' and 3' ends	
CC	CC	of the catalytic domain. Lack of expression of the hairless gene due to	
CC	CC	inherited mutations leads to the complete loss of hair, known as	
CC	CC	atrachia. Abundant hair growth, hirsutism, can be improved by inhibiting	
CC	CC	the genes promoting hair growth, and one way to get targeted, transient	
CC	CC	gene suppression is through the use of catalytic nucleic acid technology,	
CC	CC	including ribozymes and DNazymes. Ribozymes are RNA structures which have	
CC	CC	a self-catalytic enzymatic function and sequence specific RNA binding	
CC	CC	ability. Small DNA oligonucleotides that have a similar structure to the	
CC	CC	hammerhead ribozyme, called deoxy-ribozymes or DNazymes, having a	
CC	CC	catalytic core and two sequence specific arms. The deoxy-ribozymes have	



polynucleotide, polypeptide or specific binding molecule may be used to identify and detect this genetic marker in family pedigrees, radiation hybrids or human-rodent somatic cell hybrids, and as an affinity tag to identify, isolate and detect interacting proteins that bind the hairless gene or protein. This sequence represents the human hairless gene of the invention

SQ Sequence 3202 BP; 599 A; 1043 C; 1027 G; 533 T; 0 U; 0 Other;

Query Match	99.7%	Score 3190.8;	DB 8;	Length 3202;
Best Local Similarity	99.8%	Pred. No. 0;		
Matches 3195;	Conservative	0;	Mismatches	7;
			Indels	0;
			Gaps	0;

1	QY	TTTTTACTACAGGATCCGAGCATCTCCAGGTTGGCAAGAGGACCCCTTGGCAGCTGCGGAA	60
1	Db	TTTTTACTACAAGGATCCGAGCATCTCCAGGTTGGCAAGAGGACCCCTTGGCAGCTGCGGAA	60
61	QY	CTTGGGTGTTTGGCTTAAACTCTCGTGGGCACCTGACAGAGCGCGGGAGGCGGAACGC	120
61	Db	CTTGGGTGTTTGGCTTAAACTCTCGTGGGCACCTGACAGAGCGCGGGAGGCGGAACGC	120
121	QY	CTTCTACCTGCAACAGAGGATGGAGATGGGAGCTGGCCGGSCAGCAGAACTCTTGTGCCG	180
121	Db	CTTCTACCTGCAACAGAGGATGGAGATGGGAGCTGGCCGGSCAGCAGAACTCTTGTGCCG	180
181	QY	CTCTTCTCGGGCAGCAGACACTCTGCCCTTGGACCTCTGGCCGCTTGTCTCCCCAGGC	240
181	Db	CTCTTCTCGGGCAGCAGACACTCTGCCCTTGGACCTCTGGCCGCTTGTCTCCCCAGGC	240
241	QY	CTTGTTCATPACTCTTGGCAACGCTTGGCTGGGCCAGGCGATGGGAACCTTGTGGTACCAG	300
241	Db	CTTGTTCATPACTCTTGGCAACGCTTGGCTGGGCCAGGCGATGGGAACCTTGTGGTACCAG	300
301	QY	CTGGGGCCACACGACACACCCAGGTGCCCCCTCTCTGAGCGCCTGTCAACCAAGCGGGC	360
301	Db	CTGGGGCCACACGACACACCCAGGTGCCCCCTCTCTGAGCGCCTGTCAACCAAGCGGGC	360
361	QY	TGCTGTTCATCTACCCACCCACTAAAGTGGGATCTTGGCCCTTGTGGGAAAGTGGCAG	420
361	Db	TGCTGTTCATCTACCCACCCACTAAAGTGGGATCTTGGCCCTTGTGGGAAAGTGGCAG	420
421	QY	GAGGGCTGGAGGGGGTGGCAGTGGAGCCAGCGAACCCAGCGAGGAAGTGAACAAGGCC	480
421	Db	GAGGGCTGGAGGGGGTGGCAGTGGAGCCAGCGAACCCAGCGAGGAAGTGAACAAGGCC	480
481	QY	TCGTGCCCCAGGGCCCTGTCCCCCAGCCACACACCAAGCTGAAGAAGACATGGCTCACA	540
481	Db	TCGTGCCCCAGGGCCCTGTCCCCCAGCCACACACCAAGCTGAAGAAGACATGGCTCACA	540
541	QY	CGGCACTGGAGCAGTTTGAATGTCCACGCGGCTGCCCTGAGGTCGAGGAGAGGCCGGTT	600
541	Db	CGGCACTGGAGCAGTTTGAATGTCCACGCGGCTGCCCTGAGGTCGAGGAGAGGCCGGTT	600
601	QY	GCTCGGCTCCGGGCCCTCAAAGGCGCAGGACGCCGAGGTCAGGGAGCAATGGGCAT	660
601	Db	GCTCGGCTCCGGGCCCTCAAAGGCGCAGGACGCCGAGGTCAGGGAGCAATGGGCAT	660
661	QY	CCAGCCCCAAGCGGCCACCGGACCCCTTCCAGGCCCTCCAGAACAGGGGGCTGGGGGT	720
661	Db	CCAGCCCCAAGCGGCCACCGGACCCCTTCCAGGCCCTCCAGAACAGGGGGCTGGGGGT	720
721	QY	TGGCAGGAGTTTCGGGACACATCGATAGGGAACAGGATGTGACTTCGGGACAGCATGAT	780
721	Db	TGGCAGGAGTTTCGGGACACATCGATAGGGAACAGGATGTGACTTCGGGACAGCATGAT	780
781	QY	GAGCAGAAAGGACCCCAAGATGGCCAGCCAGTCTCCAGGACCCCGGACCTTCAGGACATA	840
781	Db	GAGCAGAAAGGACCCCAAGATGGCCAGCCAGTCTCCAGGACCCCGGACCTTCAGGACATA	840
841	QY	CCATGGCTGCTTCTCCCTGCAAAACTGGCTCAATGCGCAAGTTGTGCCAGCGAGCTGA	900
841	Db	CCATGGCTGCTTCTCCCTGCAAAACTGGCTCAATGCGCAAGTTGTGCCAGCGAGCTGA	900

```
QY 2161 CTTGCCCCAAAGTCAGACGAGGGCTCTGTCTCTCTGCTGCAACCGAGCTTTGGGGATGAG 2220
Db 2161 CTTGCCCCAAAGTCAGACGAGGGCTCTGTCTCTCTGCTGCAACCGAGCTTTGGGGATGAG 2220
QY 2221 GACACACGAGGGTGGAGAACCTAGCTGCCAGTCTGCCACTTCCGGAGTACTGCGCCCTC 2280
Db 2221 GACACACGAGGGTGGAGAACCTAGCTGCCAGTCTGCCACTTCCGGAGTACTGCGCCCTC 2280
QY 2281 CATGAAAACTCAACTGGCTTCTTACCTCCACCGGGCTTGCCTGGTCCACTGGAG 2340
Db 2281 CATGAAAACTCAACTGGCTTCTTACCTCCACCGGGCTTGCCTGGTCCACTGGAG 2340
QY 2341 CCCAGCTCTGGGAGCCTATGTGTGAGCCGCGACCCGGGACACTGGGGACCAAGAAC 2400
Db 2341 CCCAGCTCTGGGAGCCTATGTGTGAGCCGCGACCCGGGACACTGGGGACCAAGAAC 2400
QY 2401 CTCTGTGTGGAGTGGCCGACCTGTGTGAGCAGTCTCTGTGTCATGCCGACACACCTGCT 2460
Db 2401 CTCTGTGTGGAGTGGCCGACCTGTGTGAGCAGTCTCTGTGTCATGCCGACACACCTGCT 2460
QY 2461 GCTGTGACCGGGACAGAAAGACTTCTTTCAGGCTGGAAGGGGGTCTGGTCT 2520
Db 2461 GCTGTGACCGGGACAGAAAGACTTCTTTCAGGCTGGAAGGGGGTCTGGTCT 2520
QY 2521 CCGGCGAGCCAGTCTGTGGACGCTGTTCGGGGACAGGACGCCCGACGCGCATC 2580
Db 2521 CCGGCGAGCCAGTCTGTGGACGCTGTTCGGGGACAGGACGCCCGACGCGCATC 2580
QY 2581 CGCGCTTTCTCAGATGTGTGCCCGCGGGGACGGCCCTGTGAGCTGTGGCGCCCA 2640
Db 2581 CGCGCTTTCTCAGATGTGTGCCCGCGGGGACGGCCCTGTGAGCTGTGGCGCCCA 2640
QY 2641 GCGAGCTGCTACTGATGAGGGCTGTGGCGGGCTGTGGGAGAGTGGGGCGTGAGC 2700
Db 2641 GCGAGCTGCTACTGATGAGGGCTGTGGCGGGCTGTGGGAGAGTGGGGCGTGAGC 2700
QY 2701 TGTGTGACCTGTCTCAGACGCGCGGAGGCGGTGTGTGTGCTGCGAGGGGTCCCGAC 2760
Db 2701 TGTGTGACCTGTCTCAGACGCGCGGAGGCGGTGTGTGTGCTGCGAGGGGTCCCGAC 2760
QY 2761 CAGGTGACGGGCTGTGTGACACAGTCAGCGTCACTCAGACATTCCTCTCCCTGAGAC 2820
Db 2761 CAGGTGACGGGCTGTGTGACACAGTCAGCGTCACTCAGACATTCCTCTCCCTGAGAC 2820
QY 2821 TGTGCGCTCTCTGCTAGCTCTGCGACGAGGACCCAGCCTTCCCTGACTGCCACCTG 2880
Db 2821 TGTGCGCTCTCTGCTAGCTCTGCGACGAGGACCCAGCCTTCCCTGACTGCCACCTG 2880
QY 2881 CTTTATGCCAGATGACTGGGCTGTGTTTCAAGCAGTGAAGTGGCGGTGGGACATTA 2940
Db 2881 CTTTATGCCAGATGACTGGGCTGTGTTTCAAGCAGTGAAGTGGCGGTGGGACATTA 2940
QY 2941 CAGGAGGCCAAATAGAGGGATGTAGTGTCTGGATTCGGGTGGGGACAGGTAGACCAG 3000
Db 2941 CAGGAGGCCAAATAGAGGGATGTAGTGTCTGGATTCGGGTGGGGACAGGTAGACCAG 3000
QY 3001 GTGCTAGCCGAGCACAATTGACAGGGGATGGCGCTAGGGGACTTGGGGATTTCTGG 3060
Db 3001 GTGCTAGCCGAGCACAATTGACAGGGGATGGCGCTAGGGGACTTGGGGATTTCTGG 3060
QY 3061 TCAACCCCAAGCACCCTCTGGGACAGCAGGGGACCTCTGTTCCTTCCCTCCCTTTAG 3120
Db 3061 TCAACCCCAAGCACCCTCTGGGACAGCAGGGGACCTCTGTTCCTTCCCTCCCTTTAG 3120
QY 3121 CCAACAAACAGTGGCACCAAGTCTACCTGTCTTCTCAGGCTGGGATCTCCCGCCAC 3180
Db 3121 CCAACAAACAGTGGCACCAAGTCTACCTGTCTTCTCAGGCTGGGATCTCCCGCCAC 3180
QY 3181 CCTGTGCCCTTTTATGTACAGG 3202
Db 3181 CCTGTGCCCTTTTATGTACAGG 3202
```

RESULT 3  
ACD26351

ID ACD26351 standard; cDNA; 3202 BP.

XX ACD26351;

XX 08-SEP-2003 (first entry)

XX Human Hairless (HR) cDNA.

XX Human; hairless; HR; gene; ss; thyroid hormone receptor; hair growth;  
KW hair loss; family pedigree; radiation hybrid; somatic cell hybrid.

XX Homo sapiens.

XX Location/Qualifiers

FT 1..2955

FT /\*tag= a

FT /product= "Human hairless protein"

FT /partial

FT /transl\_except= (pos:148..150, aa:Met)

FT /transl\_except= (pos:418..420, aa:Glu)

FT /transl\_except= (pos:1150..1152, aa:Val)

FT /transl\_except= (pos:1336..1338, aa:Gln)

FT /transl\_except= (pos:1399..1401, aa:Gln)

FT /transl\_except= (pos:1585..1590, aa:Gln-Gln)

FT /transl\_except= (pos:1807..1809, aa:Lys)

FT /transl\_except= (pos:1816..1818, aa:Gln)

FT /transl\_except= (pos:1933..1935, aa:Gln)

FT /transl\_except= (pos:1975..1977, aa:Gln)

FT /transl\_except= (pos:2239..2241, aa:Asn)

FT /note= "No start codon shown"

XX US2003027300-A1.

XX 06-FEB-2003.

XX 21-DEC-2001; 2001US-00024368.

XX 07-APR-1998; 98US-0080888P.

XX 07-APR-1999; 99US-00287354.

XX (CARN-) CARNEGIE INST WASHINGTON.

XX Thompson CC;

XX WPI; 2003-492034/59.

XX P-PSDB; ABU62540.

New Hairless polypeptide and polynucleotide, useful for identifying and detecting this genetic marker in family pedigrees or human-rodent somatic cell hybrids, or detecting interacting proteins that bind hairless gene or protein.

XX Claim 3; Page 14-17; 34pp; English.

The invention relates to a human hairless (HR) polynucleotide encoding the hairless (hr) polypeptide. The invention also relates to an expression system comprising an expression construct which produces a polypeptide with hairless transcription factor activity, a reporter construct comprising a transcription regulatory region responsive to hairless transcription activity to regulate transcription of the reporter gene which is mediated by the transcription regulatory region and methods of screening for chemical agents which modulate hairless-mediated transcription, binding between hairless and thyroid hormone receptor or hairless activity. The polynucleotide is useful as a probe or primer to quantitate cognate RNA and DNA within cells, which can be subsequently used to correlate hair growth or loss with hairless expression or hairless-regulated transcription. The amino acid sequence of hairless antigen can be used for preparing specific binding molecules (e.g. polyclonal or monoclonal antibodies) for monitoring protein expression, for affinity purification and for functional studies. The human hairless



[illegible]

Qy	3061	TCACACCCCAAGCACCACCTCTGGGCACAAGCAGGGGCACTCTGTTCCCTCCCCCTTAAG	3120
Db	3061	TCACACCCCAAGCACCACCTCTGGGCACAAGCAGGGGCACTCTGTTCCCTCCCCCTTAAG	3120
Qy	3121	CCAACAACCAAGTCAGGCCACCAAGCTCACACTGTCCTTCTCAGCTGGCAATCTCCCCAC	3180
Db	3121	CCAACAACCAAGTCAGGCCACCAAGCTCACACTGTCCTTCTCAGCTGGCAATCTCCCCAC	3180
Qy	3181	CCTGTGGCCTTTATGTACAGG	3202
Db	3181	CCTGTGGCCTTTATGTACAGG	3202
RESULT 2			
AAD31077			
ID	AAD31077 standard; DNA; 3202 BP.		
XX	AC	AAD31077;	
XX	XX	18-JUN-2002 (first entry)	
XX	XX	Human Hairless gene (HR).	
XX	XX	Human; Hairless protein; Hr protein; HR gene; dermatological condition;	
KW	KW	hair loss; gene therapy; ds.	
XX	XX	Homo sapiens.	
XX	XX		
FH	Key	Location/Qualifiers	
FT	CDS	1..2955	
FT	FT	/*tag= a	
FT	FT	/product= "Human Hr protein"	
FT	FT	/transl_except= (pos:2932..2934, aa:Gly)	
FT	FT	/transl_except= (pos:2944..2946, aa:Glu)	
FT	FT	/note= "CDS does not include start codon"	
FT	FT	/partial	
XX	XX		
XX	XX	US6348348-B1.	
PN	PN		
PD	PD	19-FEB-2002.	
XX	XX		
PF	PF	07-APR-1999; 99US-00287354.	
XX	XX		
PR	PR	07-APR-1998; 98US-0080888P.	
XX	XX		
PA	PA	(CARN-) CARNEGIE INST WASHINGTON.	
XX	XX		
PI	PI	Thompson CC;	
XX	XX		
DR	DR	WPI; 2002-204622/26.	
DR	DR	P-PSDB; AAE19794.	
XX	XX		
PT	PT	Novel expression construct, useful in the diagnosis and treatment of	
PT	PT	dermatological conditions, such as hair loss, contains a Hairless gene	
PT	PT	sequence.	
XX	XX		
PS	PS	Claim 4; Col 25-32; 48pp; English.	
XX	XX		
CC	CC	The invention relates to human Hairless (Hr) polypeptides and nucleic	
CC	CC	acid molecules (HR) encoding such polypeptides. The invention also	
CC	CC	relates to the Hairless expression constructs which may be used in	
CC	CC	transcription assays. Sequences of the invention are used in diagnosis	
CC	CC	and treatment of dermatological conditions such as hair loss. They are	
CC	CC	also used in gene therapy. Polynucleotides of the invention can be used	
CC	CC	as probes for the detection of hair loss. The present sequence is human	
CC	CC	Hairless gene (HR). Note: This sequence is stated as being the same as	
CC	CC	that shown as SEQ ID NO:1 in Column 35-38 (AAD31072) of the	
CC	CC	specification. However the sequences differ at positions	
XX	XX		
XX	XX	Sequence 3202 BP; 599 A; 1044 C; 1026 G; 533 T; 0 U; 0 Other;	
Query Match 99.7%; Score 3192.4; DB 6; Length 3202;			

### Query Match

Sequence 3202 BP; 599 A; 1044 C; 1026 G; 533 T; 0 U; 0 Other; XX SO

Query Match 99.7%; Score 3192.4; DB 6; Length 3202;